Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2004/030233

International filing date: 15 September 2004 (15.09.2004)

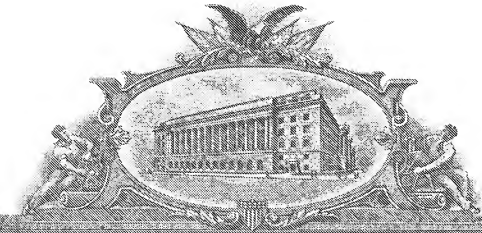
Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/504,501
Filing date: 15 September 2003 (15.09.2003)

Date of receipt at the International Bureau: 26 March 2007 (26.03.2007)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

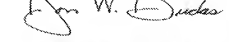
March 15, 2007

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/504,501*
FILING DATE: *September 15, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/30233*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US60/504,501*

Certified by



Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

"Express Mail" Mailing Label No. EV 351 353 889 US
Date of Deposit: **September 15, 2003**

PP-21100,001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR §1.10 on the date indicated above and is addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, Box: Provisional Application.

Nancy L. Swanson — Date: 9/15/03

PROVISIONAL APPLICATION COVER SHEET

This is a request for a PROVISIONAL APPLICATION under 37 CFR 1.53(c).

Customer No. 27476		Dkt No. PP-21100.001	Type a plus sign (+) inside this box → +
INVENTOR(S)/APPLICANT(S)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OF FOREIGN COUNTRY)
LIAN	YING		Emeryville, California
SRIVASTAVA	INDRESH	K.	Emeryville, California
BARNETT	SUSAN	W.	Emeryville, California
ROBERT-GUROFF	MARJORIE		Emeryville, California
TITLE OF INVENTION (280 characters max)			
Combination Approaches for Generating Immune Responses Against Multiple Strains Selected from a Given Subtype or Serotype			
CORRESPONDENCE ADDRESS			
CHIRON CORPORATION **Intellectual Property** **P.O. Box 8097** **Emeryville, CA 94662-8097 USA**			
STATE: **California**	ZIP CODE: **94662-8097**	COUNTRY: **USA**	
ENCLOSED APPLICATION PARTS (check all that apply)			
_X_ Specification		Number of Pages: **138**, including abstract	
_X_ Drawing(s)		Number of Pages: **28**	
___ Small Entity Statement			
_X_ Other (specify): return postcard			
METHOD OF PAYMENT (check one)			
_X_ A check or money order is enclosed to cover the Provisional filing fees			
_X_ The Commissioner is hereby authorized to charge any additional fees and credit Deposit Account Number 03-1664.		PROVISIONAL FILING FEE AMOUNT ENCLOSED: **$160.00**, CHECK NO. 8224	

The invention was made by an agency of the United States Government or under a contract with an agency of the United States government.

_X_ No

___ Yes, the name of the U.S. Government agency and the Government contract number are:

Respectfully submitted,

By: _____

Marcella Lillis
Attorney for Applicants
Registration No. 36,583

Dated: September 15, 2003

CHIRON CORPORATION
Intellectual Property
P.O. Box 8097
Emeryville, CA 94662-8097
TEL: (510) 923-8406
FAX: (510) 655-3542

PP21100.001
PATENT

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Express Mail Label No.: EV 351 353 889 US Date of Deposit: September 15, 2003

I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. § 1.10 on the date indicated above and is addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, Box: Provisional Application.

COMBINATION APPROACHES FOR GENERATING IMMUNE RESPONSES AGAINST MULTIPLE STRAINS SELECTED FROM A GIVEN SUBTYPE OR SEROTYPE

TECHNICAL FIELD

The present invention relates to compositions comprising a polynucleotide component and a polypeptide component that can be used for the generation of immune responses in a subject. In one aspect, the compositions of the present invention are used in methods to generate immune responses in subjects to which the compositions are administered. In another aspect, the compositions of the present invention are used in methods of generating neutralizing activity against multiple strains derived from a single subtype or serotype of a selected microorganims , for example, viruses (e.g., Human Immunodeficiency Virus (HIV)).

BACKGROUND OF THE INVENTION

Acquired immune deficiency syndrome (AIDS) is recognized as one of the greatest health threats facing modern medicine. There is, as yet, no cure for this disease.

In 1983-1984, three groups independently identified the suspected etiological agent of AIDS. See, e.g., Barre-Sinoussi et al. (1983) Science 220:868-871; Montagnier et al., in Human T-Cell Leukemia Viruses (Gallo, Essex & Gross, eds., 1984); Vilmer et al. (1984) The Lancet 1:753; Popovic et al. (1984) Science 224:497-500; Levy et al. (1984) Science 225:840-842. These isolates were variously called lymphadenopathy-associated virus (LAV), human T-cell lymphotropic virus type III (HTLV-III), or AIDS-associated retrovirus (ARV). All of these isolates are strains of the same virus, and were later collectively named Human Immunodeficiency Virus

(HIV). With the isolation of a related AIDS-causing virus, the strains originally called HIV are now termed HIV-1 and the related virus is called HIV-2 See, e.g., Guyader et al. (1987) Nature 326:662-669; Brun-Vezinet et al. (1986) Science 233:343-346; Clavel et al. (1986) Nature 324:691-695.

A great deal of information has been gathered about the HIV virus; however, to date an effective vaccine has not been identified. Several targets for vaccine development have been examined including the *env* and *Gag* gene products encoded by HIV. Gag gene products include, but are not limited to, Gag-polymerase and Gag-protease. Env gene products include, but are not limited to, monomeric gp120 polypeptides, oligomeric gp140 polypeptides and gp160 polypeptides.

Haas, et al., (*Current Biology* **6**(3):315-324, 1996) suggested that selective codon usage by HIV-1 appeared to account for a substantial fraction of the inefficiency of viral protein synthesis. Andre, et al., (*J. Virol.* **72**(2):1497-1503, 1998) described an increased immune response elicited by DNA vaccination employing a synthetic gp120 sequence with modified codon usage. Schneider, et al., (*J Virol.* **71(7)**:4892-4903, 1997) discuss inactivation of inhibitory (or instability) elements (INS) located within the coding sequences of the Gag and Gag-protease coding sequences.

The *Gag* proteins of HIV-1 are necessary for the assembly of virus-like particles. HIV-1 *Gag* proteins are involved in many stages of the life cycle of the virus including, assembly, virion maturation after particle release, and early post-entry steps in virus replication. The roles of HIV-1 *Gag* proteins are numerous and complex (Freed, E.O., *Virology* **251**:1-15, 1998).

Wolf, et al., (PCT International Publication No. WO 96/30523, published 3 October 1996; European Patent Application, Publication No. 0 449 116 A1, published 2 October 1991) have described the use of altered pr55 *Gag* of HIV-1 to act as a non-infectious retroviral-like particulate carrier, in particular, for the presentation of immunologically important epitopes. Wang, et al., (*Virology* **200**:524-534, 1994) describe a system to study assembly of HIV Gag-beta-galactosidase fusion proteins into virions. They describe the construction of sequences encoding HIV Gag-beta-

galactosidase fusion proteins, the expression of such sequences in the presence of HIV Gag proteins, and assembly of these proteins into virus particles.

Shiver, et al., (PCT International Publication No. WO 98/34640, published 13 August 1998) described altering HIV-1 (CAM1) *Gag* coding sequences to produce synthetic DNA molecules encoding HIV *Gag* and modifications of HIV *Gag*. The codons of the synthetic molecules were codons preferred by a projected host cell.

Recently, use of HIV Env polypeptides in immunogenic compositions has been described. (see, U.S. Patent No. 5,846,546 to Hurwitz et al., issued December 8, 1998, describing immunogenic compositions comprising a mixture of at least four different recombinant virus that each express a different HIV env variant; and U.S. Patent No. 5,840,313 to Vahlne et al., issued November 24, 1998, describing peptides which correspond to epitopes of the HIV-1 gp120 protein). In addition, U.S. Patent No. 5,876,731 to Sia et al, issued March 2, 1999 describes candidate vaccines against HIV comprising an amino acid sequence of a T-cell epitope of Gag linked directly to an amino acid sequence of a B-cell epitope of the V3 loop protein of an HIV-1 isolate containing the sequence GPGR.

PCT International Publication Nos. WO/00/39302; WO/00/39303; WO/00/39304; WO/02/04493; WO/03/004657; WO/03/004620; and WO/03/020876 described a number of codon-optimized HIV polypeptides, as well as some native HIV sequences. Further, a variety of HIV polypeptides comprising mutations were described. The use of these HIV polypeptides in vaccine compositions and methods of immunization were also described.

The present invention provides improved compositions and methods for generating immune responses against multiple subtypes, serotypes, or strains of a selected microorganism , for example, a virus (e.g., HIV-1).

SUMMARY OF THE INVENTION

The present invention relates to compositions and methods for their use for generating an immune response in a subject. The compositions of the invention comprise at least two components wherein each component provides a different but analogous polypeptide immunogen. The polypeptide immunogen is provided either

directly in the form of a polypeptide (including polypeptide fragments, modified forms, encapsulated forms, etc.) or indirectly as a polynucleotide immunogen (including DNA and/or RNA encoding a polypeptide immunogen). The compositions of the present invention may be used in methods to generate immune responses in subjects to which the compositions are administered, wherein the immune response is directed against multiple subtypes, serotypes, or strains of a selected microorganims , for example, viruses (e.g., Human Immunodeficiency Virus (HIV)). In a preferred embodiment, the present invention relates to compositions comprising a polynucleotide component and a polypeptide component that can be used for the generation of immune responses in a subject, for example, the generation of neutralizing antibodies. Other embodiments comprising at least two polynucleotide components each providing a different but analogous polypeptide immunogen, or embodiments comprising at least two polypeptide components each providing a different but analogous polypeptide immunogen are also contemplated. The compositions of the present invention may be used in methods to generate immune responses in subjects to which the compositions are administered, wherein the immune response is directed against multiple strains of a first subtype or serotype of a selected microorganims , for example, viruses (e.g., Human Immunodeficiency Virus (HIV)).

In a first aspect, the present invention includes a composition for generating an immune response in a mammal. These compositions typically comprise

a polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide derived from a first HIV strain of a first subtype, and

a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a second HIV strain of the first subtype, wherein said first HIV strain and said second HIV strain are different. A further embodiment of this composition includes the provisos that (i) the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and (ii) the polypeptide component does not comprise an

analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

In a second aspect, the present invention includes compositions for generating an immune response in a mammal. These compositions typically comprise a polynucleotide component comprising two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides derived from a first HIV subtype, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains of the first subtype, and a polypeptide component that comprises one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by the polynucleotide component, with the proviso that (i) if the polypeptide component comprises less than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then the HIV immunogenic polypeptides of the polypeptide composition may be derived from the same and/or different HIV strains of the first subtype, or (ii) if the polypeptide component comprises the same or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain of the first subtype; with the provisos that (i) the polynucleotide component does not encode an HIV immunogenic polypeptide derived from any subtype other than the first subtype, and (ii) the polypeptide component does not comprise an HIV immunogenic polypeptide derived from any subtype other than the first subtype.

The polynucleotide components of both of these aspects may comprises at least one polynucleotide that is a native polynucleotide. Alternately, or in addition, the polynucleotide components may comprise at least one polynucleotide that is a synthetic polynucleotide. Synthetic polynucleotides may comprise codons optimized for expression in mammalian cells (e.g., human cells). The polynucleotide component may comprise a single polynucleotide molecule, or two or more different polynucleotide molecules, each encoding one or more HIV polypeptides. The polynucleotide component may comprise DNA or RNA or both.

The HIV immunogenic polypeptides (encoded by the polynucleotide component and/or those which comprise the polypeptide component) may be HIV envelope polypeptides. The HIV polypeptides may comprises one or more mutations compared to the wild-type (i.e., naturally-occuring) HIV polypeptide (e.g., in the case of envelope proteins, at least one of the envelope polypeptides may comprise a mutation in the cleavage site or a mutation in the glycosylation site, a deletion or modification of the V1 region, a deletion or modification of the V2 region, a deletion or modification of the V3 region, modifications to expose an envelope binding region that binds to a CCR5 chemokine co-receptor, and combinations thereof). Other immunogenic HIV polypeptides may include, but are not limited to, Gag, Env, Pol, Prot, Int, RT, vif, vpr, vpu, tat, rev, and nef polypeptides.

The first subtype from which the HIV immunogenic polypeptides and coding sequences therefore may be selected includes, but are not limited to, the following:subtypeA,subtypeB,subtypeC,subtypeD,subtypeE,subtypeF,subtypeG, andsubtypeO, as well as any of the identified CRFs.

In addition to immunogenic HIV polypeptides and sequences encoding same, the polynucleotide component may encode and the polypeptide component may comprise one or more additional antigenic polypeptides which may include antigenic polypeptides not derived from HIV-1 coding sequences.

The polynucleotide component may further comprise sequences encoding one or more control elements compatible with expression in a selected host cell, wherein the control elements are operable linked to polynucleotides encoding HIV immunogenic polypeptides. Exemplary control elements include, but are not limited to, a transcription promoter (e.g., CMV, CMV+intron A, SV40, RSV, HIV-Ltr, MMLV-ltr, and metallothionein), a transcription enhancer element, a transcription termination signal, polyadenylation sequences, sequences for optimization of initiation of translation, internal ribosome entry sites, and translation termination sequences.

The polynucleotide component may comprise further components as described herein (e.g., carriers, vector sequences, control sequences, etc.). The polypeptide component may comprise further components as described herein (e.g., carriers, adjuvants, immunoenhancers, etc.).

The present invention also includes methods of generating an immune response in a subject. In one embodiment of the method, a composition for generating an immune response in a mammal of the present invention, for example, as described above, is provided. One or more gene delivery vectors comprising the polynucleotides of the polynucleotide component of the composition are administered to the subject under conditions that are compatible with expression of the polynucleotides in the subject for the production of encoded HIV immunogenic polypeptides. Further, the polypeptide component of the composition for generating an immune response is administered to the subject.

The one or more gene delivery vectors and the polypeptide component may be administered, for example, concurrently or sequentially.

The polynucleotide component may comprise further components as described herein (e.g., carriers, vector sequences, control sequences, etc.). The polypeptide component may comprise further components as described herein (e.g., carriers, adjuvants, immunoenhancers, etc.).

The one or more gene delivery vectors may comprise, for example, nonviral and/or viral vectors. Exemplary non-viral vectors include, but are not limited to plasmids or expression cassettes. Exemplary viral vectors include, but are not limited to retroviral, lentiviral, alphaviral, poxviral, herpes viral, adeno-associated viral, polioviral, measles viral, adenoviral vectors, or other known viral vectors. The one or more gene delivery vectors may be delivered using a particulate carrier, for example, coated on a gold or tungsten particle and the coated particle may be delivered to the subject using a gene gun, or PLG particles delivered by electroporation or otherwise. Alternatively, the one or more gene delivery vectors are encapsulated in a liposome preparation. The one or more gene delivery vectors may be administered, for example, intramuscularly, intramucosally, intranasally, subcutaneously, intradermally, transdermally, intravaginally, intrarectally, orally, intravenously, or by combinations of these methods.

The subjects of the methods of the present invention are typically mammals, for example, humans.

The immune response generated by the methods of the present invention may be humoral and/or cellular. In one embodiment, the immune response results in generating neutralizing antibodies in the subject against multiple strains derived from the first HIV subtype.

These and other embodiments of the present invention will readily occur to those of ordinary skill in the art in view of the disclosure herein.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A to 1D depict the nucleotide sequence of HIVsubtypeC 8_5_TV1_C.ZA (SEQ ID NO:1; referred to herein as TV1). Various regions are shown in Table 1.

Figures 2A-2E depicts an alignment of Env polypeptides from various HIV isolates (Type B-SF162, subtype C-TV1.8_2, subtype C-TV1.8_5, subtype C-TV2.12-5/1, subtype C-MJ4, India subtype C-93IN101, subtype A-Q2317, subtype D-92UG001, subtype E-cm235, and a Consensus Sequence). The arrows indicate exemplary regions for deletions and/or truncations in the beta and/or bridging sheet region(s). The "*" denotes N-linked glycosylation sites, one or more of which can be modified (*e.g.*, deleted and/or mutated; one such possible mutation is mutation (N→ Q)).

Figure 3 presents a schematic diagram showing the relationships between the following forms of the HIV Env polypeptide: gp160, gp140, gp120, and gp41.

Figure 4 presents neutralizing antibody activity data against HIV-1 subtype B strain SF162 obtained from a number of different immunization protocols in rabbits.

Figure 5 presents neutralizing antibody activity data against HIV-1 subtype C strain TV1 obtained from a number of different immunization protocols in rabbits.

Figure 6 presents the nucleotide sequence of the polynucleotide designated gp140.modSF162.delV2.

Figure 7 presents the nucleotide sequence of the polynucleotide designated gp140.mut7.modSF162.delV2.

Figure 8 presents the nucleotide sequence of the polynucleotide designated gp140mod.TV1.delV2.

Figure 9 presents the nucleotide sequence of the polynucleotide designated gp140mod.TV1.mut7.delV2.

Figure 10 presents the nucleotide sequence of the polynucleotide designated gp160mod.Q23-17 (optimized sequence based on subtype A HIV-1 isolate Q23-17 from Kenya GenBank Accession AF004885).

Figure 11 presents the nucleotide sequence of the polynucleotide designated gp160mod.98UA0116 (optimized sequence based on subtype A HIV-1 isolate 98UA0116 from Ukraine GenBank Accession AF413987).

Figure 12 presents the nucleotide sequence of the polynucleotide designated gp160mod.SE8538 (optimized sequence based on subtype A HIV-1 isolate SE8538 from Tanzania GenBank Accession AF069669).

Figure 13 presents the nucleotide sequence of the polynucleotide designated gp160mod.UG031 (optimized sequence based on subtype A Human immunodeficiency virus 1 proviral DNA, complete genome, clone:pUG031-A1 GenBank Accession AB098330).

Figure 14 presents the nucleotide sequence of the polynucleotide designated gp160mod.92UG001 (optimized sequence based on subtype D Human immunodeficiency virus type 1 complete proviral genome, strain 92UG001 GenBank Accession AJ320484).

Figure 15 presents the nucleotide sequence of the polynucleotide designated gp160mod.94UG114 (optimized sequence based on subtype D HIV-1 isolate 94UG114 from Uganda GenBank Accession U88824).

Figure 16 presents the nucleotide sequence of the polynucleotide designated gp160mod.ELI (optimized sequence based on subtype D Human immunodeficiency virus type 1, isolate ELIGenBank Accession K03454).

Figure 17 presents the nucleotide sequence of the polynucleotide designated gp160mod.93IN101 (optimized sequence based on Indian subtype C Human immunodeficiency virus type 1 subtype C genomic RNA GenBank Accession AB023804).

Figure 18 presents the nucleotide sequence of the polynucleotide designated gp160mod.cm235.V3con (optimized sequence based on subtype E HIV-1 isolate).

Figure 19 presents the nucleotide sequence of the polynucleotide designated gp160partialmod.cm235.V3 con (optimized sequence based on subtype E HIV-1 isolate).

Figure 20 present the ELISA data for binding antibody titers for SF162 Env protein in immunized chimpanzees.

Figure 21 presents lymphoproliferative data from chimpanzees immunized with HIV $_{MN}$ env DNA (as a prime) and HIV $_{SF162}$ env protein (as boost).

DETAILED DESCRIPTION OF THE INVENTION

The practice of the present invention will employ, unless otherwise indicated, conventional methods of chemistry, biochemistry, molecular biology, immunology and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. See, e.g., *Remington's Pharmaceutical Sciences*, 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); *Methods In Enzymology* (S. Colowick and N. Kaplan, eds., Academic Press, Inc.); and *Handbook of Experimental Immunology*, Vols. I-IV (D.M. Weir and C.C. Blackwell, eds., 1986, Blackwell Scientific Publications); Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Short Protocols in Molecular Biology*, 4th ed. (Ausubel et al. eds., 1999, John Wiley & Sons); *Molecular Biology Techniques: An Intensive Laboratory Course*, (Ream et al., eds., 1998, Academic Press); *PCR (Introduction to Biotechniques Series)*, 2nd ed. (Newton & Graham eds., 1997, Springer Verlag).

All patents, publications, sequence citations, and patent applications cited in this specification are herein incorporated by reference as if each individual patent, publication, sequence citation, or patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

As used in this specification, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more such agents.

1.0.0 DEFINITIONS

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

"Synthetic" sequences, as used herein, refers to HIV polypeptide-encoding polynucleotides whose expression has been modified as described herein, for example, by codon substitution, altered activities, and/or inactivation of inhibitory sequences. "Wild-type" or "native" sequences, as used herein, refer to polypeptide-encoding polynucleotides that are substantially as they are found in nature, e.g., Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env and/or Nef encoding sequences as found in HIV isolates, *e.g.,* SF162, SF2, AF110965, AF110967, AF110968, AF110975, MJ4 (a subtype C, Ndung'u et al. (2001) *J. Virol.* 75:4964-4972), subtype B-SF162, subtype C-TV1.8_2 (8_2_TV1_C.ZA), subtype C-TV1.8_5 (8_5_TV1_C.ZA), subtype C-TV2.12-5/1 (12-5_1_TV2_C.ZA), subtype C-MJ4, India subtype C-93IN101, subtype A-Q2317, subtype D-92UG001, subtype E-cm235, subtype A HIV-1 isolate Q23-17 from Kenya GenBank Accession AF004885, subtype A HIV-1 isolate 98UA0116 from Ukraine GenBank Accession AF413987, subtype A HIV-1 isolate SE8538 from Tanzania GenBank Accession AF069669, subtype A Human immunodeficiency virus 1 proviral DNA, complete genome, clone:pUG031-A1 GenBank Accession AB098330, subtype D Human immunodeficiency virus type 1 complete proviral genome, strain 92UG001 GenBank Accession AJ320484, subtype D HIV-1 isolate 94UG114 from Uganda GenBank Accession U88824, subtype D Human immunodeficiency virus type 1, isolate ELIGenBank Accession K03454, and Indian subtype C Human immunodeficiency virus type 1 subtype C genomic RNA GenBank Accession AB023804.

The various regions of the HIV genome are shown in Table 1, with numbering relative to 8_5_TV1_C.ZA (Figures 1-A-1D). Thus, the term "Pol" refers to one or more of the following polypeptides: polymerase (p6Pol); protease (prot); reverse transcriptase (p66RT or RT); RNAseH (p15RNAseH); and/or integrase (p31Int or Int). Identification of gene regions for any selected HIV isolate (e.g., strains within a subtype, or strains derived from different subtypes) can be performed by one of ordinary skill in the art based on the teachings presented herein and the information

known in the art, for example, by performing nucleotide and/or polypeptide alignments relative to 8_5_TV1_C.ZA (polynucleotide sequence presented in Figures 1A-1D) or alignment to other known HIV isolates, for example, Subtype B isolates with gene regions (e.g., SF2, GenBank Accession number K02007; SF162, GenBank Accession Number M38428) and Subtype C isolates with gene regions (e.g., GenBank Accession Number AF110965 and GenBank Accession Number AF110975).

As used herein, the term "virus-like particle" or "VLP" refers to a nonreplicating, viral shell, derived from any of several viruses discussed further below. VLPs are generally composed of one or more viral proteins, such as, but not limited to those proteins referred to as capsid, coat, shell, surface and/or envelope proteins, or particle-forming polypeptides derived from these proteins. VLPs can form spontaneously upon recombinant expression of the protein in an appropriate expression system. Methods for producing particular VLPs are known in the art and discussed more fully below. The presence of VLPs following recombinant expression of viral proteins can be detected using conventional techniques known in the art, such as by electron microscopy, X-ray crystallography, and the like. See, e.g., Baker et al., *Biophys. J.* (1991) 60:1445-1456; Hagensee et al., *J. Virol.* (1994) 68:4503-4505. For example, VLPs can be isolated by density gradient centrifugation and/or identified by characteristic density banding. Alternatively, cryoelectron microscopy can be performed on vitrified aqueous samples of the VLP preparation in question, and images recorded under appropriate exposure conditions.

By "particle-forming polypeptide" derived from a particular viral protein is meant a full-length or near full-length viral protein, as well as a fragment thereof, or a viral protein with internal deletions, which has the ability to form VLPs under conditions that favor VLP formation. Accordingly, the polypeptide may comprise the full-length sequence, fragments, truncated and partial sequences, as well as analogs and precursor forms of the reference molecule. The term therefore intends deletions, additions and substitutions to the sequence, so long as the polypeptide retains the ability to form a VLP. Thus, the term includes natural variations of the specified polypeptide since variations in coat proteins often occur between viral isolates. The term also includes deletions, additions and substitutions that do not naturally occur in

the reference protein, so long as the protein retains the ability to form a VLP. Preferred substitutions are those which are conservative in nature, i.e., those substitutions that take place within a family of amino acids that are related in their side chains. Specifically, amino acids are generally divided into four families: (1) acidic -- aspartate and glutamate; (2) basic -- lysine, arginine, histidine; (3) non-polar -- alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar -- glycine, asparagine, glutamine, cystine, serine threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified as aromatic amino acids.

The term "HIV polypeptide" refers to any amino acid sequence that exhibits sequence homology to native HIV polypeptides (*e.g.*, Gag, Env, Prot, Pol, RT, Int, vif, vpr, vpu, tat, rev, nef and/or combinations thereof) and/or which is functional. Non-limiting examples of functions that may be exhibited by HIV polypeptides include, use as immunogens (*e.g.*, to generate a humoral and/or cellular immune response), use in diagnostics (*e.g.* bound by suitable antibodies for use in ELISAs or other immunoassays) and/or polypeptides which exhibit one or more biological activities associated with the wild type or synthetic HIV polypeptide. For example, as used herein, the term "Gag polypeptide" may refer to a polypeptide that is bound by one or more anti-Gag antibodies; elicits a humoral and/or cellular immune response; and/or exhibits the ability to form particles.

An "antigen" refers to a molecule containing one or more epitopes (either linear, conformational or both) that will stimulate a host's immune system to make a humoral and/or cellular antigen-specific response. The term is used interchangeably with the term "immunogen." Normally, a B-cell epitope will include at least about 5 amino acids but can be as small as 3-4 amino acids. A T-cell epitope, such as a CTL epitope, will include at least about 7-9 amino acids, and a helper T-cell epitope at least about 12-20 amino acids. Normally, an epitope will include between about 7 and 15 amino acids, such as, 9, 10, 12 or 15 amino acids. The term "antigen" denotes both subunit antigens, (i.e., antigens which are separate and discrete from a whole organism with which the antigen is associated in nature), as well as, killed, attenuated or inactivated bacteria, viruses, fungi, parasites or other microbes. Antibodies such as

anti-idiotype antibodies, or fragments thereof, and synthetic peptide mimotopes, which can mimic an antigen or antigenic determinant, are also captured under the definition of antigen as used herein. Similarly, an oligonucleotide or polynucleotide which expresses an antigen or antigenic determinant *in vivo*, such as in gene therapy and DNA immunization applications, is also included in the definition of antigen herein.

For purposes of the present invention, antigens (e.g., polynucleotide encoding antigens, or polypeptides comprising antigens) can be derived from any microorganism having more than one subtype, serotype, or strain variation (e.g., viruses, bacteria, parasites, fungi, etc.). The term also intends any of the various tumor antigens. Furthermore, for purposes of the present invention, an "antigen" refers to a protein which includes modifications, such as deletions, additions and substitutions (generally conservative in nature), to the native sequence, so long as the protein maintains the ability to elicit an immunological response, as defined herein. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through mutations of hosts which produce the antigens.

An "immunological response" to an antigen or composition is the development in a subject of a humoral and/or a cellular immune response to an antigen present in the composition of interest. For purposes of the present invention, a "humoral immune response" refers to an immune response mediated by antibody molecules, while a "cellular immune response" is one mediated by T-lymphocytes and/or other white blood cells. One important aspect of cellular immunity involves an antigen-specific response by cytolytic T-cells ("CTL"s). CTLs have specificity for peptide antigens that are presented in association with proteins encoded by the major histocompatibility complex (MHC) and expressed on the surfaces of cells. CTLs help induce and promote the destruction of intracellular microbes, or the lysis of cells infected with such microbes. Another aspect of cellular immunity involves an antigen-specific response by helper T-cells. Helper T-cells act to help stimulate the function, and focus the activity of, nonspecific effector cells against cells displaying peptide antigens in association with MHC molecules on their surface. A "cellular immune response" also refers to the production of cytokines, chemokines and other

such molecules produced by activated T-cells and/or other white blood cells, including those derived from CD4+ and CD8+ T-cells.

A composition or vaccine that elicits a cellular immune response may serve to sensitize a vertebrate subject by the presentation of antigen in association with MHC molecules at the cell surface. The cell-mediated immune response is directed at, or near, cells presenting antigen at their surface. In addition, antigen-specific T-lymphocytes can be generated to allow for the future protection of an immunized host.

The ability of a particular antigen to stimulate a cell-mediated immunological response may be determined by a number of assays, such as by lymphoproliferation (lymphocyte activation) assays, CTL cytotoxic cell assays, or by assaying for T-lymphocytes specific for the antigen in a sensitized subject. Such assays are well known in the art. See, e.g., Erickson et al., *J. Immunol.* (1993) 151:4189-4199; Doe et al., *Eur. J. Immunol.* (1994) 24:2369-2376. Recent methods of measuring cell-mediated immune response include measurement of intracellular cytokines or cytokine secretion by T-cell populations, or by measurement of epitope specific T-cells (e.g., by the tetramer technique)(reviewed by McMichael, A.J., and O'Callaghan, C.A., *J. Exp. Med.* **187(9)**1367-1371, 1998; Mcheyzer-Williams, M.G., et al, *Immunol. Rev.* **150**:5-21, 1996; Lalvani, A., et al, *J. Exp. Med.* **186**:859-865, 1997).

Thus, an immunological response as used herein may be one that stimulates the production of antibodies (*e.g.,* neutralizing antibodies that block bacterial toxins and pathogens such as viruses entering cells and replicating by binding to toxins and pathogens, typically protecting cells from infection and destruction). The antigen of interest may also elicit production of CTLs. Hence, an immunological response may include one or more of the following effects: the production of antibodies by B-cells; and/or the activation of suppressor T-cells and/or memory/effector T-cells directed specifically to an antigen or antigens present in the composition or vaccine of interest. These responses may serve to neutralize infectivity, and/or mediate antibody-complement, or antibody dependent cell cytotoxicity (ADCC) to provide protection to an immunized host. Such responses can be determined using standard immunoassays and neutralization assays, well known in the art. (See, *e.g.,* Montefiori et al. (1988) *J.*

Clin Microbiol. 26:231-235; Dreyer et al. (1999) *AIDS Res Hum Retroviruses* (1999) 15(17):1563-1571).

An "immunogenic HIV polypeptide" is a polypeptide capable of elicting an immune response against one or more native HIV polypeptides, when the immunogenic polypeptide is adminstered to a laboratory test animal (such as a mouse, guina pig, rhesus macaque, chimpanzee, baboon, etc.).

An "immunogenic composition" is a composition that comprises an antigenic molecule where administration of the composition to a subject results in the development in the subject of a humoral and/or a cellular immune response to the antigenic molecule of interest. The immunogenic composition can be introduced directly into a recipient subject, such as by injection, inhalation, oral, intranasal and mucosal (*e.g.*, intra-rectally or intra-vaginally) administration.

By "subunit vaccine" is meant a vaccine composition which includes one or more selected antigens but not all antigens, derived from or homologous to, an antigen from a pathogen of interest such as from a virus, bacterium, parasite or fungus. Such a composition is substantially free of intact pathogen cells or pathogenic particles, or the lysate of such cells or particles. Thus, a "subunit vaccine" can be prepared from at least partially purified (preferably substantially purified) immunogenic polypeptides from the pathogen, or analogs thereof. The method of obtaining an antigen included in the subunit vaccine can thus include standard purification techniques, recombinant production, or synthetic production.

"Substantially purified" general refers to isolation of a substance (compound, polynucleotide, protein, polypeptide, polypeptide composition) such that the substance comprises the majority percent of the sample in which it resides. Typically in a sample a substantially purified component comprises 50%, preferably 80%-85%, more preferably 90-95% of the sample. Techniques for purifying polynucleotides and polypeptides of interest are well-known in the art and include, for example, ion-exchange chromatography, affinity chromatography and sedimentation according to density.

A "polynucleotide coding sequence" or a polynucleotide sequence that "encodes" a selected polypeptide, is a nucleic acid molecule that is transcribed (in the

case of DNA) and translated (in the case of mRNA) into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences (or "control elements"). The boundaries of the coding sequence are determined by a start codon, for example, at or near the 5' terminus and a translation stop codon, for example, at or near the 3' terminus. A coding sequence can include, but is not limited to, cDNA from viral, procaryotic or eucaryotic mRNA, genomic DNA sequences from viral or procaryotic DNA, and even synthetic DNA sequences. Exemplary coding sequences are codon optimized viral polypeptide-coding sequences used in the present invention. The coding regions of the polynucleotide sequences of the present invention are identifiable by one of skill in the art and may, for example, be easily identified by performing translations of all three frames of the polynucleotide and identifying the frame corresponding to the encoded polypeptide, for example, a synthetic nef polynucleotide of the present invention encodes a nef-derived polypeptide. A transcription termination sequence may be located 3' to the coding sequence.

Typical "control elements", include, but are not limited to, transcription regulators, such as promoters, transcription enhancer elements, transcription termination signals, and polyadenylation sequences; and translation regulators, such as sequences for optimization of initiation of translation, *e.g.*, Shine-Dalgarno (ribosome binding site) sequences, internal ribosome entry sites (IRES) such as the ECMV IRES, Kozak-type sequences (i.e., sequences for the optimization of translation, located, for example, 5' to the coding sequence, e.g., GCCACC placed in front (5') of an initiating ATG), leader sequences, translation initiation codon (e.g., ATG), and translation termination sequences (*e.g.*, TAA or, preferably, TAAA placed after (3') the coding sequence). In certain embodiments, one or more translation regulation or initiation sequences (*e.g.*, the leader sequence) are derived from wild-type translation initiation sequences, *i.e.*, sequences that regulate translation of the coding region in their native state. Wild-type leader sequences that have been modified, using the methods described herein, also find use in the present invention. Promoters can include inducible promoters (where expression of a polynucleotide sequence operably linked to the promoter is induced by an analyte, cofactor, regulatory protein, etc.), repressible promoters (where expression of a polynucleotide sequence operably linked to the

promoter is induced by an analyte, cofactor, regulatory protein, etc.), and constitutive promoters.

A "nucleic acid" molecule or "polynucleotide" can include, but is not limited to, procaryotic sequences, eucaryotic mRNA, cDNA from eucaryotic mRNA, genomic DNA sequences from eucaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. The term also captures sequences that include any of the known base analogs of DNA and RNA. In referring to the polynucleotide of the invention, in those examples in which "DNA" is specifically recited, it will be apparent that for many such embodiments, RNA is likewise intended.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, a given promoter operably linked to a coding sequence is capable of effecting the expression of the coding sequence when the proper enzymes are present. The promoter need not be contiguous with the coding sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total

DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences. Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for the gene encoding the amino acid sequence (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact amino acid to amino acid or nucleotide to nucleotide correspondence of two polypeptide sequences or polynucleotide sequences, respectively.

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison,

WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, in a preferred embodiment, BLASTN and BLASTP can be used with the following default parameters for nucleic acid searches -- genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR; (ii) polypeptide searches -- . Details of these programs can be found at the following internet address: www.ncbi.nlm.gov/cgi-bin/BLAST.

Protein similarity and percent identity sequence searches can be carried out, for example, using Smith-Waterman Similarity Search algorithms (e.g., at www.ncbi.nlm.gov, or from commercial sources, such as, TimeLogic Corporation, Crystal Bay, NV). For example, in a preferred embodiment, the Smith-Waterman Similarity Search can be used with default parameters, for example, as follows: Weight MATRIX = BLOSUM62.MAA; Gap Opening PENALTY = -12; Gap

Extension PENALTY = -2; FRAME PENALTY = 0; QUERY FORMAT = FASTA/PEARSON; QUERY TYPE = AA; QUERY SEARCH = 1; QUERY SET = CGI_1d82ws301bde.seq; TARGET TYPE = AA; TARGET SET = NRPdb gsaa; SIGNIFICANCE = GAPPED; MAX SCORES = 30; MAX ALIGNMENTS = 20; Reporting THRESHOLD = Score=1; ALIGNMENT THRESHOLD = 20.

One of skill in the art can readily determine the proper search parameters to use for a given sequence, exemplary preferred Smith Waterman based parameters are presented above. For example, the search parameters may vary based on the size of the sequence in question. Thus, for polynucleotide sequences of the present invention the length of the polynucleotide sequence disclosed herein is searched against a selected database and compared to sequences of essentially the same length to determine percent identity. For example, a representative embodiment of the present invention would include an isolated polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about a selected level of percent identity relative to Y contiguous nucleotides of one or more of the sequences described herein or fragment thereof, and (ii) for search purposes X equals Y, wherein Y is a selected reference polynucleotide of defined length (for example, a length of from 15 nucleotides up to the number of nucleotides present in a selected full-length sequence).

The sequences of the present invention can include fragments of the sequences, for example, from about 15 nucleotides up to the number of nucleotides present in the full-length sequences described herein, including all integer values falling within the above-described range. For example, fragments of the polynucleotide sequences of the present invention may be 30-60 nucleotides, 60-120 nucleotides, 120-240 nucleotides, 240-480 nucleotides, 480-1000 nucleotides, and all integer values therebetween.

The synthetic polynucleotides described herein include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% up to 100% (including all integer values falling within these described ranges) sequence identity to the synthetic polynucleotide sequences disclosed herein when the

sequences of the present invention are used as the query sequence against, for example, a database of sequences.

Two nucleic acid fragments are considered to "selectively hybridize" as described herein. The degree of sequence identity between two nucleic acid molecules affects the efficiency and strength of hybridization events between such molecules. A partially identical nucleic acid sequence will at least partially inhibit a completely identical sequence from hybridizing to a target molecule. Inhibition of hybridization of the completely identical sequence can be assessed using hybridization assays that are well known in the art (e.g., Southern blot, Northern blot, solution hybridization, or the like, see Sambrook, et al., *supra* or Ausubel et al., *supra*). Such assays can be conducted using varying degrees of selectivity, for example, using conditions varying from low to high stringency. If conditions of low stringency are employed, the absence of non-specific binding can be assessed using a secondary probe that lacks even a partial degree of sequence identity (for example, a probe having less than about 30% sequence identity with the target molecule), such that, in the absence of non-specific binding events, the secondary probe will not hybridize to the target.

When utilizing a hybridization-based detection system, a nucleic acid probe is chosen that is complementary to a target nucleic acid sequence, and then by selection of appropriate conditions the probe and the target sequence "selectively hybridize," or bind, to each other to form a hybrid molecule. A nucleic acid molecule that is capable of hybridizing selectively to a target sequence under "moderately stringent" typically hybridizes under conditions that allow detection of a target nucleic acid sequence of at least about 10-14 nucleotides in length having at least approximately 70% sequence identity with the sequence of the selected nucleic acid probe. Stringent hybridization conditions typically allow detection of target nucleic acid sequences of at least about 10-14 nucleotides in length having a sequence identity of greater than about 90-95% with the sequence of the selected nucleic acid probe. Hybridization conditions useful for probe/target hybridization where the probe and target have a specific degree of sequence identity, can be determined as is known in the art (see, for example, Nucleic Acid Hybridization: A Practical Approach, editors B.D. Hames and S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

With respect to stringency conditions for hybridization, it is well known in the art that numerous equivalent conditions can be employed to establish a particular stringency by varying, for example, the following factors: the length and nature of probe and target sequences, base composition of the various sequences, concentrations of salts and other hybridization solution components, the presence or absence of blocking agents in the hybridization solutions (e.g., formamide, dextran sulfate, and polyethylene glycol), hybridization reaction temperature and time parameters, as well as, varying wash conditions. The selection of a particular set of hybridization conditions is selected following standard methods in the art (see, for example, Sambrook, et al., *supra* or Ausubel et al., *supra*).

A first polynucleotide is "derived from" second polynucleotide if the first polynucleotide has the same basepair sequence as a region of the second polynucleotide, its cDNA, complements thereof, or if the first polynucleotide displays subtantial sequence identity to a region of the second polynucleotide, its cDNA, complements thereof, wherein sequence identity is determined as described above. Substantial sequence identity is typically about 90% or greater, preferably about 95% or greater, more preferably about 98% or greater.

A first polypeptide is "derived from" a second polypeptide if it is encoded by a first polynucleotide derived from a second polynucleotide, or the first polypeptide has the same amino acid sequence as the second polypeptide or a portion thereof, or the first polypeptide displays substantial sequence identity to the second polypeptide or a portion thereof, wherein sequence identity is determined as described above. Substantial sequence identity is typically about 90% or greater, preferably about 95% or greater, more preferably about 98% or greater.

Generally, a viral polypeptide is "derived from" a particular polypeptide of a virus (viral polypeptide) if it is (i) encoded by the same open reading frame of a polynucleotide of that virus (viral polynucleotide), or (ii) displays substantial sequence identity to a polypeptide of that virus as described above.

A polypeptide is "derived from" an HIV subtype if it is derived from a polypeptide present in a member of the subtype, derived from a polypeptide encoded by a polynucleotide present in a member of the subtype, encoded by a polynucleotide

that is derived from a polynucleotide present in a member of the subtype, or derived from a polypeptide encoded by a polynucleotide that is derived from a polynucleotide present in a member of the subtype.

A polypeptide is "derived from" an HIV strain if it is derived from a polypeptide present in a member of the strain, derived from a polypeptide encoded by a polynucleotide present in a member of the strain, encoded by a polynucleotide that is derived from a polynucleotide present in a member of the strain, or derived from a polypeptide encoded by a polynucleotide that is derived from a polynucleotide present in a member of the strain.

"Analogous polypeptides" refers to polypeptides that are encoded by, or derived from polypeptides encoded by, the same gene of the same organism but from different polynucleotide sources. In the context of the present invention, different polynucleotide sources could be different subtypes, different serotypes or different strains. Thus, for example, a Gag polypeptide from a Subtype B HIV would be an analogous polypeptide to a Gag polypeptide from a Subtype C HIV, or an envelope polypeptide derived from a first HIV-1 subtype, serotype, or strain would be an analogous polypeptide to an envelope polypeptide derived from a second HIV-1 subtype, serotype, or strain. Examples of types of analogous polypeptides that could be derived from different HIV-1 subtypes or strains include, the envelope polypeptides gp41, gp120, gp140, and gp160, all of which are considered analogous polypeptides. Further, such analogous polypeptides may each comprise different alterations or mutations, for example, analogous polypeptides derived from the HIV-1 envelope gene include, but are not limited to, the following: a gp41 polypeptide, a gp120 polypeptide, a gp140 polypeptide, a gp160 polypeptide, a gp140 comprising a deletion of a portion of the V1 loop, a gp140 polypeptide comprising a deletion of a portion of the V2 loop, a gp 140 polypeptide comprising a deletion of a portion of the V3 loop, a gp140 polypeptide with a mutated protease cleavage site, a gp160 comprising a deletion of a portion of the V1 loop, a gp160 polypeptide comprising a deletion of a portion of the V2 loop, a gp 160 polypeptide comprising a deletion of a portion of the V3 loop, and a gp160 polypeptide with a mutated protease cleavage site.

A "gene" as used in the context of the present invention is a sequence of nucleotides in a genetic nucleic acid (viral genome, chromosome, plasmid, etc.) with which a genetic function is associated. A gene is a hereditary unit, for example of an organism comprising a polynucleotide sequence (e.g., an RNA sequence for HIV-1 or a proviral HIV-1 DNA sequence), that occupies a specific physical location (a "gene locus" or "genetic locus") within the genome of an organism. A gene can encode an expressed product, such as a polypeptide or a polynucleotide (e.g., tRNA). Alternatively, a gene may define a genomic location for a particular event/function, such as the binding of proteins and/or nucleic acids (e.g., 5' LTR), wherein the gene does not encode an expressed product. Examples of HIV-1 genes include, but are not limited to, Gag, Env, Pol (prot, RNase, Int), tat, rev, nef, vif, vpr, and vpu. A gene may include coding sequences, such as, polypeptide encoding sequences, and non-coding sequences, such as, promoter sequences, poly-adenlyation sequences, transcriptional regulatory sequences (e.g., enhancer sequences). Many eucaryotic genes have "exons" (coding sequences) interrupted by "introns" (non-coding sequences). In certain cases, a gene may share sequences with another gene(s) (e.g., overlapping genes). It is noted that in the general population, wild-type genes may include multiple prevalent versions that contain alterations in sequence relative to each other. These variations are designated "polymorphisms" or "allelic variations."

"Purified polynucleotide" refers to a polynucleotide of interest or fragment thereof that is essentially free, e.g., contains less than about 50%, preferably less than about 70%, and more preferably less than about 90%, of the protein with which the polynucleotide is naturally associated. Techniques for purifying polynucleotides of interest are well-known in the art and include, for example, disruption of the cell containing the polynucleotide with a chaotropic agent and separation of the polynucleotide(s) and proteins by ion-exchange chromatography, affinity chromatography and sedimentation according to density.

By "nucleic acid immunization" is meant the introduction of a nucleic acid molecule encoding one or more selected antigens into a host cell, for the *in vivo* expression of an antigen, antigens, an epitope, or epitopes. The nucleic acid molecule can be introduced directly into a recipient subject, such as by injection, inhalation,

oral, intranasal and mucosal administration, or the like, or can be introduced *ex vivo*, into cells which have been removed from the host. In the latter case, the transformed cells are reintroduced into the subject where an immune response can be mounted against the antigen encoded by the nucleic acid molecule.

"Gene transfer" or "gene delivery" refers to methods or systems for reliably inserting nucleic acid (i.e., DNA or RNA) of interest into a host cell. Such methods can result in transient expression of non-integrated transferred DNA, extrachromosomal replication and expression of transferred replicons (e.g., episomes), or integration of transferred genetic material into the genomic DNA of host cells. Gene delivery expression vectors include, but are not limited to, vectors derived from adenoviruses, adeno-associated viruses, alphaviruses, herpes viruses, measles viruses, polio viruses, pox viruses, vesiculoviruses and vaccinia viruses. When used for immunization, such gene delivery expression vectors may be referred to as vaccines or vaccine vectors.

The term "transfection" is used to refer to the uptake of foreign DNA by a cell. A cell has been "transfected" when exogenous DNA has been introduced inside the cell membrane. A number of transfection techniques are generally known in the art. *See, e.g.*, Graham et al. (1973) *Virology*, 52:456, Sambrook et al. (1989) *Molecular Cloning, a laboratory manual*, Cold Spring Harbor Laboratories, New York, Davis et al. (1986) *Basic Methods in Molecular Biology*, Elsevier, and Chu et al. (1981) *Gene* 13:197. Such techniques can be used to introduce one or more exogenous DNA moieties into suitable host cells. The term refers to both stable and transient uptake of the genetic material, and includes uptake of peptide- or antibody-linked DNAs.

A "vector" is capable of transferring gene sequences to target cells (e.g., viral vectors, non-viral vectors, particulate carriers, and liposomes). Typically, "vector construct," "expression vector," and "gene transfer vector," mean any nucleic acid construct capable of directing the expression of a gene of interest and which can be used to transfer gene sequences to target cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors.

"Lentiviral vector", and "recombinant lentiviral vector" refer to a nucleic acid construct which carries, and within certain embodiments, is capable of directing the

expression of a nucleic acid molecule of interest. The lentiviral vector include at least one transcriptional promoter/enhancer or locus defining element(s), or other elements which control gene expression by other means such as alternate splicing, nuclear RNA export, post-translational modification of messenger, or post-transcriptional modification of protein. Such vector constructs must also include a packaging signal, long terminal repeats (LTRS) or portion thereof, and positive and negative strand primer binding sites appropriate to the retrovirus used (if these are not already present in the retroviral vector). Optionally, the recombinant lentiviral vector may also include a signal which directs polyadenylation, selectable markers such as Neo, TK, hygromycin, phleomycin, histidinol, or DHFR, as well as one or more restriction sites and a translation termination sequence. By way of example, such vectors typically include a 5' LTR, a tRNA binding site, a packaging signal, an origin of second strand DNA synthesis, and a 3'LTR or a portion thereof

"Lentiviral vector particle" as utilized within the present invention refers to a lentivirus which carries at least one gene of interest. The retrovirus may also contain a selectable marker. The recombinant lentivirus is capable of reverse transcribing its genetic material (RNA) into DNA and incorporating this genetic material into a host cell's DNA upon infection. Lentiviral vector particles may have a lentiviral envelope, a non-lentiviral envelope (e.g., an ampho or VSV-G envelope), or a chimeric envelope.

"Alphaviral vector", and "recombinant alphaviral vector" and "alphaviral replicon vector" refer to a nucleic acid construct which carries, and within certain embodiments, is capable of directing the expression of a nucleic acid molecule of interest. The alphaviral vector includes at least one transcriptional promoter/enhancer or other elements which control gene expression by other means such as alternate splicing, nuclear RNA export, post-translational modification of messenger, or post-transcriptional modification of protein. Such vector constructs must also include a packaging signal, and alphaviral replication recognition sequences. Optionally, the recombinant alphaviral vector may also include a signal which directs polyadenylation, selectable markers such as Neo, TK, hygromycin, phleomycin, histidinol, or DHFR, as well as one or more restriction sites and a translation

termination sequence. Typically, the alphaviral vector will include coding sequences for the alphaviral non-structural proteins, a packaging site, replication recognition sequences and a sequence capable of directing the expression of the nucleic acid molecule of interest.

"Expression cassette" refers to an assembly which is capable of directing the expression of a sequence or gene of interest. An expression cassette typically includes a promoter which is operably linked to the polynucleotide sequences or gene(s) of interest. Other control elements may be present as well. Expression cassettes described herein may be contained within a plasmid construct. In addition to the components of the expression cassette, the plasmid construct may also include a bacterial origin of replication, one or more selectable markers, a signal which allows the plasmid construct to exist as single-stranded DNA *(e.g.,* a M13 origin of replication), a multiple cloning site, and a "mammalian" origin of replication (e.g., a SV40 or adenovirus origin of replication).

"Packaging cell" refers to a cell that comprises those elements necessary for production of infectious recombinant viral vector, but which lack the recombinant viral vector. Typically, such packaging cells contain one or more expression cassettes that are capable of expressing proteins necessary for the replication and packaging of an introduced vector, for example, in the case of a lentiviral vector expression cassettes which encode *Gag, pol and* env proteins, in the case of an alphaviral vector, expression cassettes that encode alphaviral structural proteins.

"Producer cell" or "vector producing cell" refers to a cell which contains all elements necessary for production of recombinant viral vector particles.

Transfer of a "suicide gene" (e.g., a drug-susceptibility gene) to a target cell renders the cell sensitive to compounds or compositions that are relatively nontoxic to normal cells. Moolten, F.L. (1994) *Cancer Gene Ther.* 1:279-287. Examples of suicide genes are thymidine kinase of herpes simplex virus (HSV-tk), cytochrome P450 (Manome et al. (1996) *Gene Therapy* 3:513-520), human deoxycytidine kinase (Manome et al. (1996) *Nature Medicine* 2(5):567-573) and the bacterial enzyme cytosine deaminase (Dong et al. (1996) *Human Gene Therapy* 7:713-720). Cells which express these genes are rendered sensitive to the effects of the relatively

nontoxic prodrugs ganciclovir (HSV-tk), cyclophosphamide (cytochrome P450 2B1), cytosine arabinoside (human deoxycytidine kinase) or 5-fluorocytosine (bacterial cytosine deaminase). Culver et al. (1992) *Science* 256:1550-1552, Huber et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:8302-8306.

A "selectable marker" or "reporter marker" refers to a nucleotide sequence included in a gene transfer vector that has no therapeutic activity, but rather is included to allow for simpler preparation, manufacturing, characterization or testing of the gene transfer vector.

A "specific binding agent" refers to a member of a specific binding pair of molecules wherein one of the molecules specifically binds to the second molecule through chemical and/or physical means. One example of a specific binding agent is an antibody directed against a selected antigen.

By "subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as baboons, rhesus macaque, chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats, rabbits, and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered. The system described above is intended for use in any of the above vertebrate species, since the immune systems of all of these vertebrates operate similarly.

By "subtype" is meant a phylogenetic classification of similar organisms into groups based on similarities at the genetic (i.e., nucleic acid sequence) level. Such groups are designated "subtypes." In the HIV field, a well known and widely accepted centralized organization for the determination of such similarities and classification of particular viral isolates into subtypes is the Los Alamos National Laboratory. The HIV subtypes referred to herein are those as determined by the Los Alamos National Laboratory. (See, e.g., Myers, et al., Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico; Myers, et al., Human

Retroviruses and Aids, 1990, Los Alamos, New Mexico: Los Alamos National Laboratory.) A subtype can also be referred to as a "clade."

By "serotype" is meant a classification of similar organisms based on antibody cross-reactivity.

By "strain" is intended an organism from within the subtype but which is differentiated from other members of the same subtype based on differences in nucleic acid sequence.

By "pharmaceutically acceptable" or "pharmacologically acceptable" is meant a material which is not biologically or otherwise undesirable, i.e., the material may be administered to an individual in a formulation or composition without causing any undesirable biological effects or interacting in a deleterious manner with any of the components of the composition in which it is contained.

By "physiological pH" or a "pH in the physiological range" is meant a pH in the range of approximately 7.0 to 8.0 inclusive, more typically in the range of approximately 7.2 to 7.6 inclusive.

As used herein, "treatment" refers to any of (i) the prevention of infection or reinfection, as in a traditional vaccine, (ii) the reduction or elimination of symptoms, or (iii) the substantial or complete elimination of the pathogen in question. Treatment may be effected prophylactically (prior to infection) or therapeutically (following infection).

By "co-administration" is meant administration of more than one composition, component of a composition, or molecule. Thus, co-administration includes concurrent administration or sequentially administration (in any order), via the same or different routes of administration. Non-limiting examples of co-administration regimes include, co-administration of nucleic acid and polypeptide; co-administration of different nucleic acids (*e.g.*, different expression cassettes as described herein and/or different gene delivery vectors); and co-administration of different polypeptides (*e.g.*, different HIV polypeptides and/or different adjuvants). The term also encompasses multiple administrations of one of the co-administered molecules or compositions (*e.g.*, multiple administrations of one or more of the expression cassettes described herein followed by one or more administrations of a polypeptide-containing

composition). In cases where the molecules or compositions are delivered sequentially, the time between each administration can be readily determined by one of skill in the art in view of the teachings herein.

"T lymphocytes" or "T cells" are non-antibody producing lymphocytes that constitute a part of the cell-mediated arm of the immune system. T cells arise from immature lymphocytes that migrate from the bone marrow to the thymus, where they undergo a maturation process under the direction of thymic hormones. Here, the mature lymphocytes rapidly divide increasing to very large numbers. The maturing T cells become immunocompetent based on their ability to recognize and bind a specific antigen. Activation of immunocompetent T cells is triggered when an antigen binds to the lymphocyte's surface receptors.

2.0.0 MODES OF CARRYING OUT THE INVENTION

Before describing the present invention in detail, it is to be understood that this invention is not limited to particular formulations or process parameters as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments of the invention only, and is not intended to be limiting.

Although a number of methods and materials similar or equivalent to those described herein can be used in the practice of the present invention, the preferred materials and methods are described herein.

2.1.0 GENERAL OVERVIEW OF THE INVENTION

The present invention relates to combination approaches to generate immune responses in subjects using compositions comprising immunogenic polynucleotides and polypeptides.

In one general aspect of the present invention, a polynucleotide component of the present invention consists essentially of one polynucleotide encoding a immunogenic polypeptide derived from a microorganism (e.g., virus, bacteria, fungi, etc.), and a polypeptide component that comprises one or more immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component,

with the proviso that at least one immunogenic polypeptide of the polypeptide component is derived from a different subtype, serotype, or strain of the microorganism than the coding sequence of the immunogenic polypeptide encoded by the polynucleotide component. In this context, the polynucleotide component consisting essentially of one polynucleotide encoding an immunogenic polypeptide refers to the presence of one polynucleotide encoding one immunogenic polypeptide in the composition. The polynucleotide composition may comprise further components, such as immune enhancers, immunoregulatory components, vector sequences (e.g., viral or non-viral), carriers, particles, excipients, expression control sequences, etc. In addition, the polynucleotide component may include further components such as molecules to enhance the immune response (e.g., liposomes, PLG, particles, alum, etc.). Further, the polypeptide component may comprise further components, such as, immune enhancers, immunoregulatory components, adjuvants, carriers, particles, excipients, etc. In a further embodiment of this composition, the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and the polypeptide component does not comprise an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

In a second general aspect of the present invention, a polynucleotide component comprises two or more polynucleotide sequences comprising coding sequences for two or more analogous immunogenic polypeptides derived from a microorganism (e.g., virus, bacteria, fungi, etc.), wherein the coding sequences for at least two of the immunogenic polypeptides are derived from different subtypes, serotypes, or strains of the microorganism, and the polypeptide component comprises one or more immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that (i) if the polypeptide component provides less than the number of analogous immunogenic polypeptides encoded by the polynucleotide component, then the immunogenic polypeptides of the polypeptide composition may be derived from the same and/or different subtypes, serotypes, or strains, as the immunogenic polypeptides provided by the polynucleotide component,

or (ii) if the polypeptide component provides the same or greater than the number of analogous immunogenic polypeptides encoded by the polynucleotide component, then the immunogenic polypeptides of the polypeptide composition are derived from at least one different subtype, serotype, or strain than the immunogenic polypeptides provided by the polynucleotide component. The polynucleotide composition may comprise further components, such as immune enhancers, immunoregulatory components, vector sequences (e.g., viral or non-viral), carriers, particles, excipients, expression control sequences, etc. In addition, the polynucleotide component may include further components such as molecules to enhance the immune response (e.g., liposomes, PLG, particles, alum, etc.). Further, the polypeptide component may comprise further components, such as, immune enhancers, immunoregulatory components, adjuvants, carriers, particles, excipients, etc.

The invention is exemplified herein with reference to Human Immunodeficiency Virus 1 (HIV-1). One of ordinary skill in the art, in view of the teachings of the present specification, can apply the teachings of the present invention to other suitable organisms, for example, microorganisms. The compositions and methods of the present invention may, for example, employ polynucleotides encoding HIV envelope polypeptides and well as HIV envelope polypeptides, e.g., HIV envelope proteins analogous to those encoded by the polynucleotides, to induce broad and/or potent neutralizing activity against diverse HIV strains. Although described with reference to the HIV virus, the compositions and methods of the present invention can be applied to other virus families having a variety of subtypes, serotypes, and/or strain variations, for example, including but not limited to other non-HIV retroviruses (e.g. HTLV-1, 2), hepadnoviruses (e.g. HBV), herpesviruses (e.g. HSV-1, 2, CMV, EBV, varizella-zoster, etc.), flaviviruses (e.g. HCV, Yellow fever, Tick borne encephalitis, St. Louis Encephalitis, West Nile Virus, etc.), coronaviruses (e.g. SARS), paramyxoviruses (e.g., PIV, RSV, measles etc.), influenza viruses, picornaviruses, reoviruses (e.g., rotavirus), arenaviruses, rhabdoviruses, papovaviruses, parvoviruses, adenoviruses, Dengue virus, bunyaviruses (e.g. , hantavirus), calciviruses (e.g. Norwalk virus), filoviruses (e.g. , Ebola, Marburg).

The diversity and mutability of the HIV virus present challenges to HIV vaccine development. HIV continues to spread globally, with upwards of 42 million people infected with HIV (UNAIDS Report on the global HIV/AIDS epidemic, UNAIDS, Geneva, Switzerland (December 2002). These people are infected with different HIV subtypes (and/or strains). The infecting HIV subtype (and/or strain) is typically geographically dependent. In one aspect, the present invention relates to compositions and methods that provide the ability to induce broad and potent neutralizing antibodies against the diverse HIV subtypes, serotypes, and/or strains for the treatment of infections, reduction of infection risk, reduction of transmission, reduction of disease manifestations, and/or prevention of HIV infections arising in different regions.

Experiments performed in support of the present invention confirm the use of the combination approaches described herein to induce potent and broad HIV-neutralization activity. The approaches include immunization with a variety of polynucleotides encoding HIV polypeptides derived from different subtypes, serotypes, or strains combined with immunization using HIV polypeptides derived from different subtypes, serotypes, or strains. The invention further includes immunization using various doses and immunization regimens of such polynucleotides and polypeptides.

Accordingly, in a first particular aspect of the present invention, the polynucleotide component of the present invention consists essentially of one polynucleotide encoding an HIV immunogenic polypeptide, and the polypeptide component comprises of one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a different HIV subtype, serotype, or strain than the coding sequence of the immunogenic polypeptide encoded by the polynucleotide component. In this context, consists essentially of refers to the presence of one polynucleotide sequence encoding one HIV immunogenic polypeptide in the polynucleotide composition. The polynucleotide composition may comprise further components, such as immune enhancers, immunoregulatory components, vector sequences (e.g., viral or non-viral),

carriers, particles, excipients, expression control sequences, etc. In one embodiment of the present invention, the HIV immunogenic polypeptide encoded by the polynucleotide component is derived from subtype B, and at least one coding sequence of an HIV immunogenic polypeptide of the polypeptide component is derived from subtype C. In another embodiment, the HIV immunogenic polypeptide encoded by the polynucleotide component is derived from a first strain of a first subtype (e.g., a first subtype B strain), and at least one coding sequence of an HIV immunogenic polypeptide of the polypeptide component is derived from a second strain of the first subtype (e.g., a second subtype B strain).

In one embodiment, a polynucleotide and a polypeptide from different HIV subtypes, serotypes, or strains are used for priming and boosting, i.e., a polynucleotide encoding an immunogenic HIV polypeptide is used for immunization via delivery of the polynucleotide (e.g., a prime), an analogous immunogenic HIV polypeptide derived from a different HIV subtype, serotype, or strain is used for immunization (e.g., a boost). For example, a polynucleotide molecule is used for nucleic acid immunization, wherein the polynucleotide molecule encodes an HIV gp140 envelope polypeptide (i) derived from a South African HIV subtype C isolate/strain, (ii) that is codon optimized for expression in mammalian cells, and (iii) is mutated by deletion of the V2 loop (e.g., gp140mod.TV1.delV2, as described for example in PCT International Publication No. WO/02/04493). This nucleic acid immunization is followed by a protein boost using an HIV gp140 envelope polypeptide (i) derived from a North American HIV subtype B isolate/strain, and (ii) is mutated by deletion of the V2 loop (e.g., the protein product of gp140.mut7.modSF162.delV2, as described for example in PCT International Publication No. WO/00/39302). Oligomeric forms of the envelope polypeptide may be used (e.g., o-gp140 as described in PCT International Publication No. WO/00/39302 and US Patent No. 6,602,705). One embodiment of this aspect of the present invention, comprises a composition for generating an immune response in a mammal, the composition comprising: a polynucleotide component, comprising, a first polynucleotide encoding a first HIV immunogenic polypeptide; and a polypeptide component, comprising a second HIV immunogenic polypeptide, wherein said first and second immunogenic HIV

polypeptide are derived from different HIV subtypes, serotypes, or strains, and (ii) said first and second immunogenic polypeptides encode analogous HIV polypeptides. In one embodiment of the present invention, the analogous HIV immunogenic polypeptides coding sequences that comprise the polynucleotide composition and the HIV immunogenic polypeptides that comprise the polypeptide component of the present invention may be derived from different subtypes of HIV, in another embodiment they may derived from different strains of HIV from the same HIV subtype. In another embodiment of this aspect of the present invention the polynucleotide and polypeptide components of the present invention are used to broadly raise neutralizing antibodies against viral strains that use the CCR5 coreceptor for cell entry. For example, a composition for generating neutralizing antibodies in a mammal may comprise, a polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide derived from an HIV strain that uses the CCR5 coreceptor for cell entry, and a polypeptide component comprising one or more HIV immunogenic polypeptides derived from an HIV strain that uses the CCR5 coreceptor for cell entry analogous to the polypeptide encoded by said polynucleotide component, with the proviso that (i) if the polypeptide component has only one HIV immunogenic polypeptide, then the coding sequence of the HIV immunogenic polypeptide of the polypeptide component is derived from a different HIV strain that uses the CCR5 coreceptor for cell entry than the coding sequence of the immunogenic polypeptide encoded by the polynucleotide component, or (ii) if the polypeptide component comprises greater than one HIV immunogenic polypeptide, then the coding sequences of the polypeptides of the polypeptide component are derived from more than one HIV strain that uses the CCR5 coreceptor for cell entry.

In second particular aspect of the present invention, the polynucleotide component comprises two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV subtypes, serotypes, or strains, and the polypeptide component comprises one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that (i) if the polypeptide

component provides less than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then the HIV immunogenic polypeptides of the polypeptide composition may be derived from the same and/or different HIV subtypes, serotypes, or strains as the HIV immunogenic polypeptides provided by the polynucleotide component, or (ii) if the polypeptide component provides the same or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV subtype, serotype, or strain than the HIV immunogenic polypeptides provided by the polynucleotide component.

In one embodiment of the present invention, two or more polynucleotides encoding immunogenic HIV polypeptides, derived from at least two different subtypes, serotypes, or strains are mixed (e.g., in equal amounts) for priming. Then a single, analogous, immunogenic HIV polypeptide derived from one of the subtypes, serotypes, or strains used for priming is used for boosting. A more general embodiment comprises a composition for generating an immune response in a mammal, said composition comprising: a polynucleotide component, comprising, two or more polynucleotides each encoding analogous HIV immunogenic polypeptides, with the proviso that the coding sequences of each HIV immunogenic polypeptide are derived from different HIV subtypes, serotypes, or strains; and a polypeptide component, comprising one or more HIV immunogenic polypeptides, with the proviso that said polypeptide component comprises at least one less HIV immunogenic polypeptide than encoded by said polynucleotide component. For example, two DNA molecules are used for nucleic acid immunization, wherein the first DNA molecule encodes an HIV gp140 envelope polypeptide (i) derived from a South African HIV subtype C isolate/strain, (ii) that is codon optimized for expression in mammalian cells, and (iii) is mutated by deletion of the V2 loop (e.g., gp140mod.TV1.delV2, as described for example in PCT International Publication No. WO/02/04493), and the second DNA molecule encodes an HIV gp140 envelope polypeptide (i) derived from a North American HIV subtype B isolate, (ii) that is codon optimized for expression in mammalian cells, and (iii) is mutated by deletion of the V2 loop (e.g.,

gp140.modSF162.delV2, as described for example in PCT International Publication No. WO/00/39302). This DNA immunization is followed by a protein boost using a single HIV gp140 envelope polypeptide (i) derived from a North American HIV subtype B isolate, and (ii) is mutated by deletion of the V2 loop (e.g., the protein product of gp140.mut7.modSF162.delV2, as described for example in PCT International Publication No. WO/00/39302). Oligomeric forms of the envelope polypeptide may be used (e.g., o-gp140 as described in PCT International Publication No. WO/00/39302). One embodiment of a composition for generating an immune response in a mammal comprises, a polynucleotide component comprising a first polynucleotide encoding a first immunogenic HIV polypeptide, and a second polynucleotide encoding a second immunogenic HIV polypeptide, wherein (i) said first and second immunogenic HIV polypeptide are derived from different HIV subtypes, serotypes, or strains, and (ii) said first and second immunogenic polypeptides encode analogous HIV polypeptides, and a polypeptide component comprising said first HIV immunogenic polypeptide, or said second HIV immunogenic polypeptide, with the proviso that said polypeptide component comprises at least one less HIV immunogenic polypeptide than is encoded by said polynucleotide component. In a preferred embodiment, polynucleotides encoding analogous immunogenic HIV polypeptides, derived from a variety of different HIV subtypes, serotypes, or strains are used for a prime immunization, and a single analogous immunogenic HIV polypeptide is used for one or more protein boost.

In another embodiment, two or more polynucleotides encoding immunogenic HIV polypeptides, derived from at least two different subtypes, serotypes, or strains are mixed (e.g., in equal amounts) for priming. Then one or more analogous, immunogenic HIV polypeptides derived from at least two different subtypes, serotypes, or strains are used for boosting, wherein at least one of the immunogenic HIV polypeptides is derived from a subtype, serotype, or strain not represented in the polynucleotide component. For example, the polynucleotide component comprises three polynucleotides encoding three immunogenic HIV polypeptides, one coding sequence derived from a subtype B strain, one coding sequence derived from a subtype C strain, and one coding sequence derived from a subtype E strain, and the

polypeptide component comprises three immunogenic HIV polypeptides, one coding sequence derived from a subtype B strain, one coding sequence derived from a subtype C strain, and one coding sequence derived from a subtype O strain. In another embodiment of this aspect of the present invention, the polynucleotides of the polynucleotide component comprises polynucleotides encoding analogous HIV immunogenic polypeptides from different subtypes, serotypes, or strains as the polypeptides of the polypeptide component. For example, DNA immunization with two or more DNA molecules encoding HIV gp140 polypeptides (wherein the two or more gp140 coding sequences are derived from two or more HIV-1 subtypes, serotypes, or strains). The polypeptide component, use for protein immunization, comprises two or more gp140 polypeptides (wherein the two or more gp140 coding sequences are derived from two or more HIV-1 subtypes, serotypes, or strains, with the proviso that at least one of the polypeptide sequences is derived from an HIV-1 subtype, serotype, or strain not represented in the DNA component).

In another embodiment, the polynucleotide component comprises two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains that use the CCR5 coreceptor for cell entry, and the polypeptide component comprises one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that (i) if the polypeptide component provides less than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then the HIV immunogenic polypeptides of the polypeptide composition may be derived from the same and/or different HIV strains that use the CCR5 coreceptor for cell entry as the HIV immunogenic polypeptides provided by the polynucleotide component, or (ii) if the polypeptide component provides the same or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain that uses the CCR5 coreceptor for cell entry than the HIV immunogenic polypeptides provided by the polynucleotide component.

In a further aspect, the present invention relates to the use of varied doses of polynucleotides and polypeptides in prime/boost methods, particularly the methods described herein. In any immunization method using, for example, a mixed polynucleotide prime (i.e., two or more polynucleotides encoding immunogenic HIV polypeptides derived from two or more HIV subtypes, serotypes, or strains) in conjunction with a polypeptide boost the present invention includes using reduced doses of each single component to provide an equivalent immune response to using full doses of each component. In one embodiment, the high threshold of DNA is the maximum tolerable dose of DNA (e.g., about 5 mg to about 10 mg total DNA), the low threshold of DNA is the minimum effective dose (e.g., about 2 ug to about 10 ug total DNA), the high threshold of protein is the maximum tolerable dose of protein (e.g., about 1 mg total protein), the low threshold of protein is the minimum effective dose (e.g., about 2 ug total protein). Experiments performed in support of the present invention demonstrated that the total DNA dose may be divided among the polynucleotides of the polynucleotide component (for example, four polynucleotide constructs used, the total DNA for all four is less than or equal to the high threshold) (e.g., Example 4). Further, the total polypeptide dose may be divided among the polypeptides comprising the polypeptide component (for example, four polypeptides used, the total protein for all four is less than or equal to the high threshold) (e.g., Example 4). The total DNA and total protein are both typically above the low threshold values.

In a preferred embodiment, the total amount of DNA in a given DNA immunization has a high threshold of less than or equal to about 10 mg total DNA and greater than or equal to 1 mg total DNA, and the total amount of protein in a given polypeptide boost has a high threshold of less than or equal to about 200 ug total protein product and greater than or equal to 10 ug of total protein. For example, in an embodiment using a polynucleotide component having two DNA molecules each encoding an immunogenic HIV polypeptide the dose of each DNA molecule per subject may be one milligram of each DNA molecule encoding an immunogenic HIV polypeptide, for a total of 2 mg for the two DNA molecules, or 0.5 mg of each DNA molecule encoding an immunogenic HIV polypeptide, for a total of 1 mg for the two

DNA molecules. Dosing with the polypeptide component may be similarly varied, for example, using a polypeptide component having two immunogenic HIV polypeptides the dose of each polypeptide per subject may be 100 micrograms of each immunogenic HIV polypeptide, for a total of 200 ug for the two polypeptides, 50 micrograms of each immunogenic HIV polypeptide, for a total of 100 ug for the two polypeptides, or 25 ug of each immunogenic HIV polypeptide, for a total of 50 ug for the two polypeptides. As described above, more than two polypeptides may be included in the polypeptide component of the present invention.

In one embodiment of this aspect of the present invention, the polynucleotides of the polynucleotide component encode analogous HIV immunogenic polypeptides from the same subtypes, serotypes, or strains as the polypeptides of the polypeptide component. For example, two DNA molecules are used for nucleic acid immunization, wherein the first DNA molecule encodes an HIV gp140 envelope polypeptide (i) derived from a South African HIV subtype C isolate, (ii) that is codon optimized for expression in mammalian cells, (iii) is mutated by deletion of the V2 loop (e.g., gp140mod.TV1.delV2, as described for example in PCT International Publication No. WO/02/04493), and (iv) is delivered at 0.5 mg, and the second DNA molecule encodes an HIV gp140 envelope polypeptide (i) derived from a North American HIV subtype B isolate, (ii) that is codon optimized for expression in mammalian cells, (iii) is mutated by deletion of the V2 loop (e.g., gp140.modSF162.delV2, as described for example in PCT International Publication No. WO/00/39302), and (iv) is delivered at 0.5 mg. This DNA immunization is followed by a protein boost using an HIV gp140 envelope polypeptide (i) derived from a South African HIV subtype C isolate, (ii) is mutated by deletion of the V2 loop (e.g., the protein product of gp140mod.TV1.mut7.delV2, as described for example in PCT International Publication No. WO/02/04493), and (iii) is delivered at 50 ug protein, and an HIV gp140 envelope polypeptide (i) derived from a North American HIV subtype B isolate, (ii) is mutated by deletion of the V2 loop (e.g., the protein product of gp140.mut7.modSF162.delV2, as described for example in PCT International Publication No. WO/00/39302), and (iii) is delivered at 50 ug protein.

Further, oligomeric forms of the envelope polypeptide may be used (e.g., o-gp140 as described in PCT International Publication No. WO/00/39302).

In further embodiments, the polynucleotide component of the present invention may comprise one or more gene delivery vectors comprising the polynucleotide(s) encoding immunogenic HIV polypeptide(s). Further components that may be included in the polynucleotide component are described herein. The polypeptide component of the present invention may comprise an adjuvant in addition to the immunogenic polypeptide(s). Further components that may be included in the polypeptide component are described herein.

The present invention also comprises methods for generating an immune response in a subject. In one general aspect, the method comprises administering to a subject a first component providing an immunogenic polypeptide and administering to a subject a second component providing a different but analogous immunogenic polypeptide. The first component and the second component may be polynucleotide components or polypeptide components. The immunogenic polypeptides may be provided either directly (as in a polypeptide component) or indirectly (as in a polynucleotide component). In a preferred embodiment, one of the components (either first or second component) is a polynucleotide component, and the other component (either second or first component) is a polypeptide component. Preferably, the polypeptide immunogens provided by the first and second components are analogous HIV immunogenic polypeptides. The first and second components may be administered simultaneously or may be administered at separate times. Preferably, the first and second components are administered in a prime-boost regimen. Various prime-boost regimens have been described in the art and are well known to those of ordinary skill. In a typical prime-boost regimen, a first component providing a polypeptide immunogen is administered to a subject; the initial immune response is followed by determining the production of binding antibodies to the polypeptide immunogen in said subject until the titer of binding antibodies begins to decline; and a second component providing a different but related polypeptide immunogen is administered to the subject.

The first and second components may be provided as a composition. In a particular aspect the method comprises, providing a composition of the present invention for generating an immune response in a mammal, administering one or more gene delivery vectors comprising the polynucleotides of the polynucleotide component of the composition into the subject under conditions that are compatible with expression of the polynucleotides in the subject for the production of encoded HIV immunogenic polypeptides, and administering the polypeptide component to the subject. The administering of the polynucleotide and polypeptide compositions may be concurrent or sequentially. In a preferred embodiment immunization with the polynucleotide component precedes immunization with the polypeptide component. Further, a single prime may be followed by multiple boosts, multiple primes may be followed by a single boost, multiple primes may be followed by multiple boosts, or a series of primes and boosts may be used. The polynucleotide component may comprise further components (e.g., components for enhancing immune response, carriers, etc.). The polypeptide component may comprise further components (e.g., components for enhancing immune response, carriers, etc.).

Exemplary polynucleotide constructs, methods of making the polynucleotide constructs, corresponding polypeptide products, and methods of making polypeptides useful for HIV immunization have been previously described, for example, in the following PCT International Publication Nos.: WO/00/39302; WO/00/39303; WO/00/39304; WO/02/04493; WO/03/004657; WO/03/004620; and WO/03/020876.

Although described generally with reference to HIV subtypes B and C as exemplary subtypes, the compositions and methods of the present invention are applicable to a wide variety of HIV subtypes, serotypes, or strains and immunogenic polypeptides encoded thereby, including but not limited to the previously identified HIV-1 subtypes A through K, N and O, the identified CRFs (circulating recombinant forms), and HIV-2 strains and its subtypes. See, e.g., Myers, et al., Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico; Myers, et al., Human Retroviruses and Aids, 1990, Los Alamos, New Mexico: Los Alamos National Laboratory. Further, the compositions and methods of the present invention may be used to raise broadly reactive neutralizing antibodies against viral strains and subtypes

that use the CCR5 coreceptor for cell entry (for example, both TV1 and SF162 use the CCR5 coreceptor (Example 4)).

The polypeptide component of the present invention may comprise fragments of immunogenic polypeptide, for example, wherein the polypeptide sequence or a portion thereof contains an amino acid sequence of at least 3 to 5 amino acids, more preferably at least 8 to 10 amino acids, and even more preferably at least 15 to 20 amino acids from a polypeptide encoded by the nucleic acid sequence. Also encompassed are polypeptide sequences that are immunologically identifiable with a polypeptide encoded by the sequence. Further, polyproteins can be constructed by fusing in-frame two or more polynucleotide sequences encoding polypeptide or peptide products.

In addition, the polynucleotide component of the present invention may comprise one or more monocistronic expression cassettes comprising polynucleotides encoding immunogenic HIV polypeptides, or one or more polycistronic expression cassettes comprising polynucleotides encoding immunogenic HIV polypeptides, or combinations thereof. Polycistronic coding sequences may be produced, for example, by placing two or more polynucleotide sequences encoding polypeptide products adjacent each other, typically under the control of one promoter, wherein each polypeptide coding sequence may be modified to include sequences for internal ribosome binding sites.

A variety of combinations of polynucleotides encoding immunogenic polypeptides (e.g., HIV immunogenic polypeptides) and immunogenic polypeptides or fragments thereof (e.g., HIV immunogenic polypeptides) can be used in the practice of the present invention. Polynucleotide sequences encoding immunogenic polypeptides can be included in a polynucleotide component of compositions of the present invention, for example, as DNA immunization constructs containing, for example, a synthetic Env expression cassettes, a synthetic Gag expression cassette, a synthetic pol-derived polypeptide expression cassette, a synthetic expression cassette comprising sequences encoding one or more accessory or regulatory genes (e.g., tat, rev, nef, vif, vpu, vpr). Immunogenic polypeptides may be included as purified polypeptides in the polypeptide component of compositions of the present invention.

The immunogenic polypeptides may be synthetic or wild-type. In preferred embodiments the immunogenic polypeptides are antigenic viral proteins, or fragments thereof.

2.2.0 IDENTIFICATION OF ANALOGOUS POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SUCH POLYPEPTIDES

The compositions and methods of the present invention are described with reference to exemplary HIV-1 sequences. The present invention is not limited to the sequences described herein. Numerous sequences for use in the practice of the present invention have been previously described (see, e.g., PCT International Publication Nos. WO/00/39302; WO/00/39303; WO/00/39304; WO/02/04493; WO/03/004657; WO/03/004620; and WO/03/020876.). Typically, the polynucleotide sequences used in the practice of the present invention encode polypeptides derived from a viral source (e.g., HIV-1). The polypeptides are typically derived from antigenic viral proteins, in particular, group specific antigen polypeptides, envelope polypeptides, capsid polypeptides, and other structural and non-structural polypeptides. The present invention is particularly described with reference to the use of envelope polypeptides and modifications thereof (and polynucleotides encoding same) derived from various subtypes, serotypes, or strains of the HIV-1 virus. Other HIV-1 polypeptides and polynucleotides encoding such polypeptides may be used in the practice of the present invention including, but not limited to, Gag, Pol (including Protease, Reverse Transcriptase, and Integrase), Tat, Rev, Nef, Vif, Vpr, and Vpu.

The HIV genome and various polypeptide-encoding regions are shown in Table 1. The nucleotide positions are given relative to an HIV-1 Subtype C isolate from South Africa strain 8_5_TV1_C.ZA (Figures 1A-1D). However, it will be readily apparent to one of ordinary skill in the art in view of the teachings of the present disclosure how to determine corresponding regions in other HIV strains (from the same or different subtypes) or variants (*e.g.*, isolates HIV_{IIIb}, HIV_{SF2}, $HIV\text{-}1_{SF162}$, $HIV\text{-}1_{SF170}$, HIV_{LAV}, HIV_{LAI}, HIV_{MN}, $HIV\text{-}1_{CM235}$,, $HIV\text{-}1_{US4}$, other HIV-1 strains from diverse subtypes(e.g., subtypes, A through K, N and O), the identified CRFs (circulating recombinant forms), HIV-2 strains and diverse subtypes and strains (e.g.,

HIV-2_{UC1} and HIV-2_{UC2}), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); *Virology*, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (*e.g.*, BLAST and others described herein) or identification and alignment of structural features (*e.g.*, a program such as the "ALB" program described herein that can identify the various regions).

Table 1

Regions of the HIV Genome relative to the Sequence of 8_5_TV1_C.ZA

Region	Position in nucleotide sequence
5'LTR	**1-636**
U3	1-457
R	458-553
U5	554-636
NFkB II	340-348
NFkB I	354-362
Sp1 III	379-388
Sp1 II	390-398
Sp1 I	400-410
TATA Box	429-433
TAR	474-499
Poly A signal	529-534
PBS	**638-655**
p7 binding region, packaging signal	**685-791**
Gag:	**792-2285**
p17	792-1178

Region	Position in nucleotide sequence
p24	1179-1871
Cyclophilin A bdg.	1395-1505
MHR	1632-1694
p2	1872-1907
p7	1908-2072
Frameshift slip	2072-2078
p1	2073-2120
p6Gag	2121-2285
Zn-motif I	1950-1991
Zn-motif II	2013-2054
Pol:	**2072-5086**
p6Pol	2072-2245
Prot	2246-2542
p66RT	2543-4210
p15RNaseH	3857-4210
p31Int	4211-5086
Vif:	**5034-5612**
Hydrophilic region	5292-5315
Vpr:	**5552-5839**
Oligomerization	5552-5677
Amphipathic a-helix	5597-5653
Tat:	**5823-6038 and 8417-8509**
Tat-1 exon	5823-6038
Tat-2 exon	8417-8509
N-terminal domain	5823-5885
Trans-activation domain	5886-5933

Region	Position in nucleotide sequence
Transduction domain	5961-5993
Rev:	**5962-6037 and 8416-8663**
Rev-1 exon	5962-6037
Rev-2 exon	8416-8663
High-affinity bdg. site	8439-8486
Leu-rich effector domain	8562-8588
Vpu:	**6060-6326**
Transmembrane domain	6060-6161
Cytoplasmic domain	6162-6326
Env (gp160):	**6244-8853**
Signal peptide	6244-6324
gp120	6325-7794
V1	6628-6729
V2	6727-6852
V3	7150-7254
V4	7411-7506
V5	7663-7674
C1	6325-6627
C2	6853-7149
C3	7255-7410
C4	7507-7662
C5	7675-7794
CD4 binding	7540-7566
gp41	7795-8853
Fusion peptide	7789-7842
Oligomerization domain	7924-7959
N-terminal heptad repeat	7921-8028

Region	Position in nucleotide sequence
C-terminal heptad repeat	8173-8280
Immunodominant region	8023-8076
Nef:	**8855-9478**
Myristoylation	8858-8875
SH3 binding	9062-9091
Polypurine tract	9128-9154
SH3 binding	9296-9307

It will be readily apparent that one of skill in the art can align any HIV sequence to that shown in Table 1 to determine relative locations of any particular HIV gene. For example, using one of the alignment programs described herein (*e.g.*, BLAST), other HIV genomic sequences can be aligned with 8_5_TV1_C.ZA (Table 1) and locations of genes determined. Polypeptide sequences can be similarly aligned. For example, Figures 2A-2E shows the alignment of Env polypeptide sequences from various strains, relative to SF-162. As described in detail in PCT International Publication No. WO/00/39303, Env polypeptides (*e.g.*, gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel beta-strands (beta-2, beta-3, beta -20 and beta -21) that form a beta -sheet. Extruding from one pair of the beta -strands (beta -2 and beta -3) are two loops, V1 and V2. The beta -2 sheet occurs at approximately amino acid residue 113 (Cys) to amino acid residue 117 (Thr) while beta -3 occurs at approximately amino acid residue 192 (Ser) to amino acid residue 194 (Ile), relative to SF-162. The "V1/V2 region" occurs at approximately amino acid positions 120 (Cys) to residue 189 (Cys), relative to SF-162. Extruding from the second pair of beta -strands (beta -20 and beta -21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." The locations of both the small loop and bridging sheet small loop can be determined relative to HXB-2 following the teachings herein and in PCT International Publication No. WO/00/39303. Also shown by arrows in Figures 2A-2E are approximate sites for deletions sequence from the beta sheet region. The "*" denotes N-glycosylation sites that can be mutated following the teachings of the present specification.

2.3.0 Expression Cassettes Comprising Polynucleotide Sequences, Vectors, Polypeptides, Further Components, and Formulations Useful in the Practice of the Present Invention

Compositions for the generation of immune responses of the present invention comprise a polynucleotide component and a polypeptide component. The polynucleotide component of may comprise one or more polynucleotides encoding immunogenic viral polypeptides. Such polynucleotides may comprise native viral sequences encoding immunogenic viral polypeptides or synthetic polynucleotides encoding immunogenic polypeptides. Synthetic polynucleotides may include sequence optimization to provide improved expression of the encoded polypeptides relative to the analogous native polynucleotide sequences. Further, synthetic polynucleotides may comprise mutations (single or multiple point mutations, missense mutations, nonsense mutations, deletions, insertions, etc.) relative to corresponding wild-type sequences.

The polypeptide component of the compositions of the present invention may comprise one or more immunogenic viral polypeptide. Such polypeptides may comprise native immunogenic viral polypeptides or modified immunogenic polypeptides. Modified polypeptides may include sequence optimization to provide improved expression of the polypeptides relative to the analogous native polynucleotide sequences. Further, modified polypeptides may comprise mutations (single or multiple point mutations, missense mutations, nonsense mutations, deletions, insertions, etc.) relative to corresponding wild-type sequences.

The compositions of the present invention, comprising a polynucleotide component and a polypeptide component, are described with reference to HIV-1 derived sequences. However, the compositions and methods of the present invention are applicable to other types of viruses as well, wherein such viruses comprise multiple subtypes, serotypes, and/or strain variations, for example, including but not limited to other non-HIV retroviruses (e.g. HTLV-1, 2), hepadnoviruses (e.g. HBV), herpesviruses (e.g. HSV-1, 2, CMV, EBV, varizella-zoster, etc.), flaviviruses (e.g. HCV, Yellow fever, Tick borne encephalitis, St. Louis Encephalitis, West Nile Virus,

etc.), coronaviruses (e.g. SARS), paramyxoviruses (e.g., PIV, RSV, measles etc.), influenza viruses, picornaviruses, reoviruses (e.g., rotavirus), arenaviruses, rhabdoviruses, papovaviruses, parvoviruses, adenoviruses, Dengue virus, bunyaviruses (e.g. , hantavirus), calciviruses (e.g. Norwalk virus), filoviruses (e.g. , Ebola, Marburg).

2.3.1 Modification of Polynucleotide Coding Sequences

HIV-1 coding sequences, and related sequences, may be modified to have improved expression in target cells relative to the corresponding wild-type sequences. Following here are some exemplary modifications that can be made to such coding sequences.

First, the HIV-1 codon usage pattern may be modified so that the resulting nucleic acid coding sequence are comparable to codon usage found in highly expressed human genes. The HIV codon usage reflects a high content of the nucleotides A or T of the codon-triplet. The effect of the HIV-1 codon usage is a high AT content in the DNA sequence that results in a decreased translation ability and instability of the mRNA. In comparison, highly expressed human codons prefer the nucleotides G or C. The HIV coding sequences may be modified to be comparable to codon usage found in highly expressed human genes.

Second, there are inhibitory (or instability) elements (INS) located within the coding sequences of, for example, the Gag coding sequences. The RRE is a secondary RNA structure that interacts with the HIV encoded Rev-protein to overcome the expression down-regulating effects of the INS. To overcome the post-transcriptional activating mechanisms of RRE and Rev, the instability elements can be inactivated by introducing multiple point mutations that do not alter the reading frame of the encoded proteins.

Third, for some genes the coding sequence has been altered such that the polynucleotide coding sequence encodes a gene product that is inactive or non-functional (e.g., inactivated polymerase, protease, tat, rev, nef, vif, vpr, and/or vpu gene products). Example 1 describes some exemplary mutations.

The synthetic coding sequences are assembled by methods known in the art, for example by companies such as the Midland Certified Reagent Company (Midland, Texas), following the guidance of the present specification.

Some exemplary synthetic polynucleotide sequences encoding immunogenic HIV polypeptides and the polypeptides encoded thereby for use in the methods of the present invention have been described, for example, in PCT International Publication Nos. WO/00/39303, WO/00/39302, WO 00/39304, WO/02/04493, WO/03/020876, WO/03/004620, and WO/03/004657.

In a preferred embodiment, the present invention relates to polynucleotides encoding Env polypeptides and corresponding Env polypeptides. For example, the codon usage pattern for Env may be modified so that the resulting nucleic acid coding sequence is comparable to codon usage found in highly expressed human genes. Such synthetic Env sequences are capable of higher level of protein production relative to the native Env sequences (see, for example, PCT International Publication Nos. WO/00/39302). Modification of the Env polypeptide coding sequences results in improved expression relative to the wild-type coding sequences in a number of mammalian cell lines (as well as other types of cell lines, including, but not limited to, insect cells). Similar Env polypeptide coding sequences can be obtained, modified and tested for improved expression from a variety of isolates.

Further modifications of Env include, but are not limited to, generating polynucleotides that encode Env polypeptides having mutations and/or deletions therein. For instance, the hypervariable regions, V1 and/or V2, can be deleted as described herein. In addition, the variable regions V3, V4 and/ or V5 can be modified or deleted. (See e.g, US 6,602,705) Additionally, other modifications, for example to the bridging sheet region and/or to N-glycosylation sites within Env can also be performed following the teachings of the present specification. (see, Figures 2A-2E, as well as PCT International Publication Nos. WO/00/39303, WO/00/39302, WO 00/39304, WO/02/04493, WO/03/020876, and WO/03/004620). Other useful modifications of env are well known and include those described in Schulke et al., (J. Virol. 2002 76:7760), Yang et al. 2002, (J. Virol. 2002 76:4634), Yang et al. 2001(J. Virol. 2001 75:1165), Shu et al. (Biochem. 1999 38:5378), Farzan et al. (J.Virol. 1998

72:7620) and Xiang et al. (J.Virol. 2002 76:9888). Various combinations of these modifications can be employed to generate synthetic expression cassettes and corresponding polypeptides as described herein.

The present invention also includes expression cassettes which include synthetic sequences derived HIV genes other than Env, including but not limited to, regions within Gag, Env, Pol, as well as, tat, rev, nef, vif, vpr, and vpu. Further, the present invention includes synthetic polynucleotides and/or expression cassettes (as well as polypeptide encoded thereby) comprising two or more antigenic polypeptides. Such sequences may be used, for example, in their entirety or sequences encoding specific epitopes or antigens may be selected from the synthetic coding sequences following the teachings of the present specification and information known in the art. For example, the polypeptide sequences encoded by the polynucleotides may be subjected to computer analysis to predict antigenic peptide fragments within the full-length sequences. The corresponding polynucleotide coding fragments may then be used in the constructs of the present invention. Exemplary algorithms useful for such analysis include, but are not limited to, the following:

AMPHI. This program has been used to predict T-cell epitopes (Gao, et al., (1989) J. Immunol. **143**:3007; Roberts, et al, (1996) AIDS Res Hum Retrovir **12**:593; Quakyi, et al., (1992) Scand J Immunol suppl. **11**:9). The AMPHI algorithm is available int the Protean package of DNASTAR, Inc. (Madison, WI, USA).

ANTIGENIC INDEX. This algorithm is useful for predicting antigenic determinants (Jameson & Wolf, (1998) CABIOS 4:181:186; Sherman, KE, et al., Hepatology 1996 Apr;23(4):688-94; Kasturi, KN, et al, J Exp Med 1995 Mar 1;181(3):1027-36; van Kampen V, et al., Mol Immunol 1994 Oct;31(15):1133-40; Ferroni P, et al., J Clin Microbiol 1993 Jun;31(6):1586-91; Beattie J, et al., Eur J Biochem 1992 Nov 15;210(1):59-66; Jones GL, et al, Mol Biochem Parasitol 1991 Sep;48(1):1-9).

HYDROPHILICITY. One algorithm useful for determining antigenic determinants from amino acid sequences was disclosed by Hopp & Woods (1981) (PNAS USA **78:**3824-3828.

Default parameters, for the above-recited algorithms, may be used to determine antigenic sites. Further, the results of two or more of the above analyses may be combined to identify particularly preferred fragments.

2.3.2 Further Modification of Polynucleotide Sequences And Polypeptides Encoded Thereby

The immunogenic viral polypeptide-encoding expression cassettes described herein may also contain one or more further sequences encoding, for example, one or more transgenes. In one embodiment of the present invention, the polynucleotide component may comprise coding sequences for one or more HIV immunogenic polypeptides. Further, the polypeptide component may comprise one or more HIV immunogenic polypeptide. In a different embodiment of the present invention, a polynucleotide component may comprise coding sequences for one or more HIV immunogenic polypeptides, wherein the polynucleotide component further comprises a sequence encoding an additional antigenic polypeptide, with the proviso that the additional antigenic polypeptide is not an immunogenic polypeptide derived from an HIV-1 strain. Further, the polypeptide component may comprise one or more HIV immunogenic polypeptides, wherein the polypeptide component further comprises an additional antigenic polypeptide, with the proviso that the additional antigenic polypeptide is not an immunogenic polypeptide derived from an HIV-1 strain.

Further sequences (*e.g.*, transgenes) useful in the practice of the present invention include, but are not limited to, further sequences are those encoding further viral epitopes/antigens {including but not limited to, HCV antigens (e.g., E1, E2; Houghton, M.., et al., U.S. Patent No. 5,714,596, issued February 3, 1998; Houghton, M.., et al., U.S. Patent No. 5,712,088, issued January 27, 1998; Houghton, M.., et al., U.S. Patent No. 5,683,864, issued November 4, 1997; Weiner, A.J., et al., U.S. Patent No. 5,728,520, issued March 17, 1998; Weiner, A.J., et al., U.S. Patent No. 5,766,845, issued June 16, 1998; Weiner, A.J., et al., U.S. Patent No. 5,670,152, issued September 23, 1997), HIV antigens (e.g., derived from one or more HIV isolate); and sequences encoding tumor antigens/epitopes. Further sequences may also be derived from non-viral sources, for instance, sequences encoding cytokines such interleukin-2

(IL-2), stem cell factor (SCF), interleukin 3 (IL-3), interleukin 6 (IL-6), interleukin 12 (IL-12), G-CSF, granulocyte macrophage-colony stimulating factor (GM-CSF), interleukin-1 alpha (IL-1alpha), interleukin-11 (IL-11), MIP-1, tumor necrosis factor (TNF), leukemia inhibitory factor (LIF), c-kit ligand, thrombopoietin (TPO) and flt3 ligand, commercially available from several vendors such as, for example, Genzyme (Framingham, MA), Genentech (South San Francisco, CA), Amgen (Thousand Oaks, CA), R&D Systems and Immunex (Seattle, WA). Additional sequences are described herein below.

HIV polypeptide coding sequences can be obtained from other HIV isolates, see, e.g., Myers et al. Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico (1992); Myers et al., *Human Retroviruses and Aids, 1997*, Los Alamos, New Mexico: Los Alamos National Laboratory. Synthetic expression cassettes can be generated using such coding sequences as starting material by following the teachings of the present specification.

Further, the synthetic expression cassettes of the present invention include related polypeptide sequences having greater than 85%, preferably greater than 90%, more preferably greater than 95%, and most preferably greater than 98% sequence identity to the polypeptides encoded by the synthetic expression cassette sequences disclosed herein.

Exemplary expression cassettes and modifications are set forth in Example 1 and are discussed further herein below.

Further, the polynucleotides of the present invention may comprise alternative polymer backbone structures such as, but not limited to, polyvinyl backbones (Pitha, *Biochem Biophys Acta*, 204:39, 1970a; Pitha, *Biopolymers*, 9:965, 1970b), and morpholino backbones (Summerton, J., *et al.*, U.S. Patent No. 5,142,047, issued 08/25/92; Summerton, J., *et al.*, U.S. Patent No. 5,185,444 issued 02/09/93). A variety of other charged and uncharged polynucleotide analogs have been reported. Numerous backbone modifications are known in the art, including, but not limited to, uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoamidates, and carbamates) and charged linkages (*e.g.*, phosphorothioates and phosphorodithioates.

2.3.3 Exemplary Cloning Vectors and Systems for Use with the Polynucleotide Sequences Encoding Immunogenic Polypeptides

Polynucleotide sequences for use in the compositions and methods of the present invention can be obtained using recombinant methods, such as by screening cDNA and genomic libraries from cells expressing the gene, or by deriving the gene from a vector known to include the same. Furthermore, the desired gene can be isolated directly from cells and tissues containing the same, using standard techniques, such as phenol extraction and PCR of cDNA or genomic DNA. See, e.g., Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA. The gene of interest can also be produced synthetically, rather than cloned. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. In general, one will select preferred codons for the intended host in which the sequence will be expressed. The complete sequence is assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge, *Nature* (1981) 292:756; Nambair et al., *Science* (1984) 223:1299; Jay et al., *J. Biol. Chem.* (1984) 259:6311; Stemmer, W.P.C., (1995) *Gene* **164**:49-53.

Next, the gene sequence encoding the desired antigen can be inserted into a vector containing a synthetic expression cassette of the present invention. In one embodiment, polynucleotides encoding selected antigens are separately cloned into expression vectors (e.g., a first Env-coding polynucleotide in a first vector, a second analogous Env-coding polynucleotide in a second vector). In certain embodiments, the antigen is inserted into or adjacent a synthetic Gag coding sequence such that when the combined sequence is expressed it results in the production of VLPs comprising the Gag polypeptide and the antigen of interest, *e.g.,* Env (native or modified) or other antigen(s) (native or modified) derived from HIV. Insertions can be made within the coding sequence or at either end of the coding sequence (5', amino terminus of the expressed Gag polypeptide; or 3', carboxy terminus of the expressed Gag polypeptide)(Wagner, R., et al., *Arch Virol.* **127**:117-137, 1992; Wagner, R., et al., *Virology* **200**:162-175, 1994; Wu, X., et al., *J. Virol.* **69(6)**:3389-3398, 1995;

Wang, C-T., et al., *Virology* **200**:524-534, 1994; Chazal, N., et al., *Virology* **68(1)**:111-122, 1994; Griffiths, J.C., et al., *J. Virol.* **67(6)**:3191-3198, 1993; Reicin, A.S., et al., *J. Virol.* **69(2)**:642-650, 1995). Up to 50% of the coding sequences of p55Gag can be deleted without affecting the assembly to virus-like particles and expression efficiency (Borsetti, A., et al, *J. Virol.* **72(11)**:9313-9317, 1998; Gamier, L., et al., *J Virol* **72(6)**:4667-4677, 1998; Zhang, Y., et al., *J Virol* **72(3)**:1782-1789, 1998; Wang, C., et al., *J Virol* **72(10)**: 7950-7959, 1998). When sequences are added to the amino terminal end of Gag, the polynucleotide can contain coding sequences at the 5' end that encode a signal for addition of a myristic moiety to the Gag-containing polypeptide (e.g., sequences that encode Met-Gly).

Expression cassettes for use in the practice of the present invention can also include control elements operably linked to the coding sequence that allow for the expression of the gene *in vivo* in the subject species. For example, typical promoters for mammalian cell expression include the SV40 early promoter, a CMV promoter such as the CMV immediate early promoter, the mouse mammary tumor virus LTR promoter, the adenovirus major late promoter (Ad MLP), and the herpes simplex virus promoter, among others. Other nonviral promoters, such as a promoter derived from the murine metallothionein gene, will also find use for mammalian expression. Typically, transcription termination and polyadenylation sequences will also be present, located 3' to the translation stop codon. Preferably, a sequence for optimization of initiation of translation, located 5' to the coding sequence, is also present. Examples of transcription terminator/polyadenylation signals include those derived from SV40, as described in Sambrook et al., *supra*, as well as a bovine growth hormone terminator sequence.

Enhancer elements may also be used herein to increase expression levels of the mammalian constructs. Examples include the SV40 early gene enhancer, as described in Dijkema et al., *EMBO J.* (1985) 4:761, the enhancer/promoter derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus, as described in Gorman et al., *Proc. Natl. Acad. Sci. USA* (1982b) 79:6777 and elements derived from human CMV, as described in Boshart et al., *Cell* (1985) 41:521, such as elements included in the CMV intron A sequence.

Furthermore, plasmids can be constructed which include a chimeric antigen-coding gene sequences, encoding, e.g., multiple antigens/epitopes of interest, for example derived from more than one viral isolate.

Typically the antigen coding sequences precede or follow the synthetic coding sequence and the chimeric transcription unit will have a single open reading frame encoding both the antigen of interest and the synthetic coding sequences. Alternatively, multi-cistronic cassettes (e.g., bi-cistronic cassettes) can be constructed allowing expression of multiple antigens from a single mRNA using the EMCV IRES, or the like.

In one embodiment of the present invention, the polynucleotide component of an immune generating composition may comprise, for example, the following: a first expression vector comprising a first Env expression cassette, wherein the Env coding sequence is derived from a first HIV subtype, serotype, or strain, and a second expression vector comprising a second Env expression cassette, wherein the Env coding sequence is derived from a second HIV subtype, serotype, or strain. Expression cassettes comprising coding sequences of the present invention may be combined in any number of combinations depending on the coding sequence products (e.g., HIV polypeptides) to which, for example, an immunological response is desired to be raised. In yet another embodiment, synthetic coding sequences for multiple HIV-derived polypeptides may be constructed into a polycistronic message under the control of a single promoter wherein IRES are placed adjacent the coding sequence for each encoded polypeptide.

Exemplary polynucleotide sequences of interest for use in the present invention may be derived from strains including, but not limited to: subtype B-SF162, subtype C-TV1.8_2 (8_2_TV1_C.ZA), subtype C-TV1.8_5 (8_5_TV1_C.ZA), subtype C-TV2.12-5/1 (12-5_1_TV2_C.ZA), subtype C-MJ4, India subtype C-93IN101, subtype A-Q2317, subtype D-92UG001, subtype E-cm235, subtype A HIV-1 isolate Q23-17 from Kenya GenBank Accession AF004885, subtype A HIV-1 isolate 98UA0116 from Ukraine GenBank Accession AF413987, subtype A HIV-1 isolate SE8538 from Tanzania GenBank Accession AF069669, subtype A Human immunodeficiency virus 1 proviral DNA, complete genome, clone:pUG031-A1

GenBank Accession AB098330, subtype D Human immunodeficiency virus type 1 complete proviral genome, strain 92UG001 GenBank Accession AJ320484, subtype D HIV-1 isolate 94UG114 from Uganda GenBank Accession U88824, subtype D Human immunodeficiency virus type 1, isolate ELIGenBank Accession K03454, and Indian subtype C Human immunodeficiency virus type 1 subtype C genomic RNA GenBank Accession AB023804.

Polynucleotide coding sequences used in the present invention may encode functional gene products or be mutated to reduce (relative to wild-type), attenuate, inactivate, eliminate, or render non-functional the activity of the gene product(s) encoded the synthetic polynucleotide.

Once complete, the expression cassettes are typically used in constructs for nucleic acid immunization using standard gene delivery protocols. Methods for gene delivery are known in the art. See, e.g., U.S. Patent Nos. 5,399,346, 5,580,859, 5,589,466. Genes can be delivered either directly to the vertebrate subject or, alternatively, delivered *ex vivo*, to cells derived from the subject and the cells reimplanted in the subject.

A number of viral based systems have been developed for gene transfer into mammalian cells. Selected sequences can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to cells of the subject either *in vivo* or *ex vivo*. A number of viral based systems have been developed for use as gene transfer vectors for mammalian host cells. For example, retroviruses (in particular, lentiviral vectors) provide a convenient platform for gene delivery systems. A coding sequence of interest (for example, a sequence useful for gene therapy applications) can be inserted into a gene delivery vector and packaged in retroviral particles using techniques known in the art. Recombinant virus can then be isolated and delivered to cells of the subject either *in vivo* or *ex vivo*. A number of retroviral systems have been described, including, for example, the following: (U.S. Patent No. 5,219,740; Miller et al. (1989) *BioTechniques* 7:980; Miller, A.D. (1990) *Human Gene Therapy* 1:5; Scarpa et al. (1991) *Virology* 180:849; Burns et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033; Boris-Lawrie et al. (1993) *Cur. Opin. Genet. Develop.* 3:102; GB 2200651; EP

0415731; EP 0345242; PCT International Publication No. WO 89/02468; PCT International Publication No. WO 89/05349; PCT International Publication No. WO 89/09271; PCT International Publication No. WO 90/02806; PCT International Publication No. WO 90/07936; PCT International Publication No. WO 90/07936; PCT International Publication No. WO 94/03622; PCT International Publication No. WO 93/25698; PCT International Publication No. WO 93/25234; PCT International Publication No. WO 93/11230; PCT International Publication No. WO 93/10218; PCT International Publication No. WO 91/02805; in U.S. 5,219,740; U.S. 4,405,712; U.S. 4,861,719; U.S. 4,980,289 and U.S. 4,777,127; in U.S. Serial No. 07/800,921; and in Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53:83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci USA* 81;6349; and Miller (1990) *Human Gene Therapy* 1.

One type of retrovirus, the murine leukemia virus, or "MLV", has been widely utilized for gene therapy applications (see generally Mann et al. *(Cell 33:153,* 1993), Cane and Mulligan *(Proc, Nat'l. Acad. Sci. USA* 81:6349, 1984), and Miller et al., *Human Gene Therapy* 1:5-14,1990.

Lentiviral vectors may be readily constructed from a wide variety of lentiviruses *(see* RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985). Representative examples of lentiviruses included HIV, HIV-1, HIV-2, FIV and SIV. Such lentiviruses may either be obtained from patient isolates, or, more preferably, from depositories or collections such as the American Type Culture Collection, or isolated from known sources using available techniques. Portions of the lentiviral gene delivery vectors (or vehicles) may be derived from different viruses. For example, in a given recombinant lentiviral vector, LTRs may be derived from an HIV, a packaging signal from SIV, and an origin of second strand synthesis from HrV-2. Lentiviral vector constructs may comprise a 5' lentiviral LTR, a tRNA binding site, a packaging signal, one or more heterologous sequences, an origin of second strand DNA synthesis and a 3' LTR. The lentiviral vectors have a nuclear transport element that, in preferred embodiments is not RRE. Representative examples of suitable nuclear transport elements include the element in Rous sarcoma

virus (Ogert, et al., J *ViroL 70,* 3834-3843, 1996), the element in Rous sarcoma virus (Liu & *Mertz, Genes & Dev., 9,* 1766-1789, 1995) and the element in the genome of simian retrovirus type 1 (Zolotukhin, et al., *J Virol. 68,* 7944-7952, 1994). Other potential elements include the elements in the histone gene (Kedes, *Annu. Rev. Biochem. 48,* 837-870, 1970), interferon gene (Nagata et al., *Nature 287,* 401-408, 1980), adrenergic receptor gene (Koilka, et al., *Nature 329,* 75-79, 1987), and the c-Jun gene (Hattorie, et al., *Proc. Natl. Acad. Sci. USA 85,* 9148-9152, 1988).

A number of adenovirus vectors have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham, *J. Virol.* (1986) 57:267-274; Bett et al., *J. Virol.* (1993) 67:5911-5921; Mittereder et al., *Human Gene Therapy* (1994) 5:717-729; Seth et al., *J. Virol.* (1994) 68:933-940; Barr et al., *Gene Therapy* (1994) 1:51-58; Berkner, K.L. *BioTechniques* (1988) 6:616-629; and Rich et al., *Human Gene Therapy* (1993) 4:461-476).

Additionally, various adeno-associated virus (AAV) vector systems have been developed for gene delivery. AAV vectors can be readily constructed using techniques well known in the art. See, e.g., U.S. Patent Nos. 5,173,414 and 5,139,941; PCT International Publication Nos. WO 92/01070 (published 23 January 1992) and WO 93/03769 (published 4 March 1993); Lebkowski et al., *Molec. Cell. Biol.* (1988) 8:3988-3996; Vincent et al., *Vaccines 90* (1990) (Cold Spring Harbor Laboratory Press); Carter, B.J. *Current Opinion in Biotechnology* (1992) 3:533-539; Muzyczka, N. *Current Topics in Microbiol. and Immunol.* (1992) 158:97-129; Kotin, R.M. *Human Gene Therapy* (1994) 5:793-801; Shelling and Smith, *Gene Therapy* (1994) 1:165-169; and Zhou et al., *J. Exp. Med.* (1994) 179:1867-1875.

Another vector system useful for delivering the polynucleotides of the present invention is the enterically administered recombinant poxvirus vaccines described by Small, Jr., P.A., et al. (U.S. Patent No. 5,676,950, issued October 14, 1997).

Additional viral vectors that will find use for delivering the nucleic acid molecules encoding the antigens of interest include those derived from the pox family of viruses, including vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the genes can be constructed as follows. The DNA

encoding the particular immunogenic HIV polypeptide coding sequence is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells that are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the coding sequences of interest into the viral genome. The resulting TK recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the genes. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an avipox vector is particularly desirable in human and other mammalian species since members of the avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, e.g., PCT International Publication Nos. WO 91/12882; WO 89/03429; and WO 92/03545.

Molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al., *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al., *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery.

Members of the Alphavirus genus, such as, but not limited to, vectors derived from the Sindbis, Semliki Forest, and Venezuelan Equine Encephalitis viruses, will also find use as viral vectors for delivering the polynucleotides of the present invention (for example, first and second synthetic gp140-polypeptide encoding expression cassette, wherein the first and second gp140 polypeptides are analogous and derived from different HIV subtypes, serotypes, or strains). For a description of Sindbis-virus derived vectors useful for the practice of the instant methods, see, Dubensky et al., *J. Virol.* (1996) 70:508-519; and PCT International Publication Nos. WO 95/07995 and WO 96/17072; as well as, Dubensky, Jr., T.W., et al., U.S. Patent No. 5,843,723, issued December 1, 1998, and Dubensky, Jr., T.W., U.S. Patent No.

5,789,245, issued August 4, 1998. Preferred expression systems include, but are not limited to, eucaryotic layered vector initiation systems (e.g., US Patent No. 6,015,686, US Patent No. 5, 814,482, US Patent No. 6,015,694, US Patent No. 5,789,245, EP 1029068A2, PCT International Publication No. WO 9918226A2/A3, EP 00907746A2, PCT International Publication No. WO 9738087A2). Exemplary expression systems include, but are not limited to, chimeric alphavirus replicon particles, for example, those that form VEE and SIN (see, e.g., Perri, et al., "An alphavirus replicon particle chimera derived from Venezuelan equine encephalitis and Sindbis viruses is a potent gene-based vaccine delivery vector," J. Virol 2003, 77(19), in press; PCT WO02/099035; USSN 10/310734, filed Dec 4 2002). Such alphavirus-based vector systems can be used in a prime or as a boost in DNA-primed subjects or potentially as a stand-alone immunization method for the induction of neutralizing antibodies using the multivalent approaches described herein.

Expression cassette delivery vectors may also include tissue-specific promoters to drive expression of one or more genes or sequences of interest.

Expression cassette delivery vector constructs may be generated such that more than one gene of interest is expressed. This may be accomplished through the use of di- or oligo-cistronic cassettes (e.g., where the coding regions are separated by 80 nucleotides or less, *see generally* Levin et al., *Gene* 108:167-174, 1991), or through the use of Internal Ribosome Entry Sites ("IRES").

Synthetic expression cassettes of interest can also be delivered without a viral vector. For example, delivery of the expression cassettes of the present invention can also be accomplished using eucaryotic expression vectors comprising CMV-derived elements, such vectors include, but are not limited to, the following: pCMVKm2, pCMV-link pCMVPLEdhfr, and pCMV6a (see Example 1). For example, a synthetic DNA expression cassette of the present invention, e.g., one encoding gp140 polypeptide, may be cloned into the following eucaryotic expression vectors: pCMVKm2, for transient expression assays and DNA immunization studies, the pCMVKm2 vector is derived from pCMV6a (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986) and comprises a kanamycin selectable marker, a ColE1 origin of replication, a CMV promoter enhancer and Intron A, followed by an insertion site for

the synthetic sequences described below followed by a polyadenylation signal derived from bovine growth hormone -- the pCMVKm2 vector differs from the pCMV-link vector only in that a polylinker site is inserted into pCMVKm2 to generate pCMV-link; pESN2dhfr and pCMVPLEdhfr, for expression in Chinese Hamster Ovary (CHO) cells; and, pAcC13, a shuttle vector for use in the Baculovirus expression system (pAcC13, is derived from pAcC12 which is described by Munemitsu S., et al., *Mol Cell Biol.* **10(11)**:5977-5982, 1990).

In additon, the expression cassettes of the present invention can be packaged in liposomes prior to delivery to the subject or to cells derived therefrom. Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed DNA to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight, *Biochim. Biophys. Acta.* (1991) 1097:1-17; Straubinger et al., in *Methods of Enzymology* (1983), Vol. 101, pp. 512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations, with cationic liposomes particularly preferred. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner et al., *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416); mRNA (Malone et al., *Proc. Natl. Acad. Sci. USA* (1989) 86:6077-6081); and purified transcription factors (Debs et al., *J. Biol. Chem.* (1990) 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner et al., *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416). Other commercially available lipids include (DDAB/DOPE) and DOTAP/DOPE (Boerhinger). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g., Szoka et al., *Proc. Natl. Acad. Sci. USA* (1978) 75:4194-4198; PCT International Publication No. WO 90/11092 for a description of

the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as, from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilammelar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See, e.g., Straubinger et al., in METHODS OF IMMUNOLOGY (1983), Vol. 101, pp. 512-527; Szóka et al., *Proc. Natl. Acad. Sci. USA* (1978) 75:4194-4198; Papahadjopoulos et al., *Biochim. Biophys. Acta* (1975) 394:483; Wilson et al., *Cell* (1979) 17:77); Deamer and Bangham, *Biochim. Biophys. Acta* (1976) 443:629; Ostro et al., *Biochem. Biophys. Res. Commun.* (1977) 76:836; Fraley et al., *Proc. Natl. Acad. Sci. USA* (1979) 76:3348); Enoch and Strittmatter, *Proc. Natl. Acad. Sci. USA* (1979) 76:145); Fraley et al., *J. Biol. Chem.* (1980) 255:10431; Szoka and Papahadjopoulos, *Proc. Natl. Acad. Sci. USA* (1978) 75:145; and Schaefer-Ridder et al., *Science* (1982) 215:166.

The DNA and/or protein antigen(s) can also be delivered in cochleate lipid compositions similar to those described by Papahadjopoulos et al., *Biochem. Biophys. Acta.* (1975) 394:483-491. See, also, U.S. Patent Nos. 4,663,161 and 4,871,488.

The expression cassettes of interest may also be encapsulated, adsorbed to, or associated with, particulate carriers. Such carriers present multiple copies of a selected antigen to the immune system and promote trapping and retention of antigens in local lymph nodes. The particles can be phagocytosed by macrophages and can enhance antigen presentation through cytokine release. Examples of particulate carriers include those derived from polymethyl methacrylate polymers, as well as

microparticles derived from poly(lactides) and poly(lactide-co-glycolides), known as PLG. See, e.g., Jeffery et al., *Pharm. Res.* (1993) 10:362-368; McGee JP, et al., *J Microencapsul.* 14(2):197-210, 1997; O'Hagan DT, et al., *Vaccine* 11(2):149-54, 1993. Suitable microparticles may also be manufactured in the presence of charged detergents, such as anionic or cationic detergents, to yield microparticles with a surface having a net negative or a net positive charge. For example, microparticles manufactured with anionic detergents, such as hexadecyltrimethylammonium bromide (CTAB), i.e. CTAB-PLG microparticles, adsorb negatively charged macromolecules, such as DNA. (see, e.g., Int'l Application Number PCT/US99/17308).

Furthermore, other particulate systems and polymers can be used for the *in vivo* or *ex vivo* delivery of the gene of interest. For example, polymers such as polylysine, polyarginine, polyornithine, spermine, spermidine, as well as conjugates of these molecules, are useful for transferring a nucleic acid of interest. Similarly, DEAE dextran-mediated transfection, calcium phosphate precipitation or precipitation using other insoluble inorganic salts, such as strontium phosphate, aluminum silicates including bentonite and kaolin, chromic oxide, magnesium silicate, talc, and the like, will find use with the present methods. See, e.g., Felgner, P.L., *Advanced Drug Delivery Reviews* (1990) 5:163-187, for a review of delivery systems useful for gene transfer. Peptoids (Zuckerman, R.N., et al., U.S. Patent No. 5,831,005, issued November 3, 1998) may also be used for delivery of a construct of the present invention.

In some embodiments of the present invention, alum and PLG are useful delivery adjuvants that enhance immunity to polynucleotide vaccines (e.g., DNA vaccines). Further embodiments include, but are not limited to, toxoids, cytokines, and co-stimulatory molecules may also be used as genetic adjuvants with polynucleotide vaccines.

Additionally, biolistic delivery systems employing particulate carriers such as gold and tungsten, are especially useful for delivering synthetic expression cassettes of the present invention. The particles are coated with the synthetic expression cassette(s) to be delivered and accelerated to high velocity, generally under a reduced atmosphere, using a gun powder discharge from a "gene gun." For a description of

such techniques, and apparatuses useful therefore, see, e.g., U.S. Patent Nos. 4,945,050; 5,036,006; 5,100,792; 5,179,022; 5,371,015; and 5,478,744. Also, needle-less injection systems can be used (Davis, H.L., et al, *Vaccine* 12:1503-1509, 1994; Bioject, Inc., Portland, OR).

Recombinant vectors carrying a synthetic expression cassette of the present invention are formulated into compositions for delivery to the vertebrate subject. These compositions may either be prophylactic (to prevent infection) or therapeutic (to treat disease after infection). If prevention of disease is desired, the compositions are generally administered prior to primary infection with the pathogen of interest. If treatment is desired, e.g., the reduction of symptoms or recurrences, the compositions are generally administered subsequent to primary infection. The compositions will comprise a "therapeutically effective amount" of the gene of interest such that an amount of the antigen can be produced *in vivo* so that an immune response is generated in the individual to which it is administered. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the subject to be treated; the capacity of the subject's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular antigen selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. Thus, a "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

The compositions will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, polyethyleneglycol, hyaluronic acid, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Certain facilitators of nucleic acid uptake and/or expression can also be included in the compositions or coadministered, such as, but not limited to, bupivacaine, cardiotoxin and sucrose.

Once formulated, the compositions of the invention can be administered directly to the subject (e.g., as described above) or, alternatively, delivered *ex vivo*, to cells derived from the subject, using methods such as those described above. For

example, methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and can include, e.g., dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, lipofectamine and LT-1 mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) (with or without the corresponding antigen) in liposomes, and direct microinjection of the DNA into nuclei.

Direct delivery of synthetic expression cassette compositions *in vivo* will generally be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). The constructs can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of polypeptides encoding immunogenic polypeptides is combined with administration of analogous immunogenic polypeptides following the methods of the present invention.

2.3.4 Expression of Synthetic Sequences Encoding HIV-1 Polypeptides and Related Polypeptides

Immunogenic viral polypeptide-encoding sequences of the present invention can be cloned into a number of different expression vectors/host cell systems to provide immunogenic polypeptides for the polypeptide component of the immune-response generating compositions of the present invention. For example, DNA fragments encoding HIV polypeptides can be cloned into eucaryotic expression vectors, including, a transient expression vector, CMV-promoter-based mammalian vectors, and a shuttle vector for use in baculovirus expression systems. Synthetic polynucleotide sequences (e.g., codon optimized polynucleotide sequences) and wild-

type sequences can typically be cloned into the same vectors. Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. See, generally, Sambrook et al, *supra*. The vector is then used to transform an appropriate host cell. Suitable recombinant expression systems include, but are not limited to, bacterial, mammalian, baculovirus/insect, vaccinia, Semliki Forest virus (SFV), Alphaviruses (such as, Sindbis, Venezuelan Equine Encephalitis (VEE)), mammalian, yeast and Xenopus expression systems, well known in the art. Particularly preferred expression systems are mammalian cell lines, vaccinia, Sindbis, eucaryotic layered vector initiation systems (e.g., US Patent No. 6,015,686, US Patent No. 5, 814,482, US Patent No. 6,015,694, US Patent No. 5,789,245, EP 1029068A2, PCT International Publication No. WO 9918226A2/A3, EP 00907746A2, PCT International Publication No. WO 9738087A2), insect and yeast systems.

A number of host cells for such expression systems are also known in the art. For example, mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (A.T.C.C.), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find use with the present expression constructs. Yeast hosts useful in the present invention include *inter alia*, *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guillerimondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Insect cells for use with baculovirus expression vectors include, *inter alia*, *Aedes aegypti, Autographa californica, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda*, and *Trichoplusia ni*. See, e.g., Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987).

Viral vectors can be used for expression of polypeptides in eucaryotic cells, such as those derived from the pox family of viruses, including vaccinia virus and avian poxvirus. For example, a vaccinia based infection/transfection system, as described in Tomei et al., *J. Virol.* (1993) 67:4017-4026 and Selby et al., *J. Gen. Virol.*

(1993) 74:1103-1113, will also find use with the present invention. A vaccinia based infection/transfection system can be conveniently used to provide for inducible, transient expression of the coding sequences of interest in a host cell. In this system, cells are first infected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA that is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, e.g., Elroy-Stein and Moss, *Proc. Natl. Acad. Sci. USA* (1990) 87:6743-6747; Fuerst et al., *Proc. Natl. Acad. Sci. USA* (1986) 83:8122-8126.

As an alternative approach to infection with vaccinia or avipox virus recombinants, an amplification system can be used that will lead to high level expression following introduction into host cells. Specifically, a T7 RNA polymerase promoter preceding the coding region for T7 RNA polymerase can be engineered. Translation of RNA derived from this template will generate T7 RNA polymerase which in turn will transcribe more template. Concomitantly, there will be a cDNA whose expression is under the control of the T7 promoter. Thus, some of the T7 RNA polymerase generated from translation of the amplification template RNA will lead to transcription of the desired gene. Because some T7 RNA polymerase is required to initiate the amplification, T7 RNA polymerase can be introduced into cells along with the template(s) to prime the transcription reaction. The polymerase can be introduced as a protein or on a plasmid encoding the RNA polymerase. For a further discussion of T7 systems and their use for transforming cells, see, e.g., PCT International Publication No. WO 94/26911; Studier and Moffatt, *J. Mol. Biol.* (1986) 189:113-130; Deng and Wolff, *Gene* (1994) 143:245-249; Gao et al., *Biochem. Biophys. Res. Commun.* (1994) 200:1201-1206; Gao and Huang, *Nuc. Acids Res.* (1993) 21:2867-2872; Chen et al., *Nuc. Acids Res.* (1994) 22:2114-2120; and U.S. Patent No. 5,135,855.

These vectors are transfected into an appropriate host cell. The cell lines are then cultured under appropriate conditions and the levels of any appropriate polypeptide product can be evaluated in supernatants. For example, p24 can be used to evaluate Gag expression; gp160, gp140 or gp120 can be used to evaluate Env expression; p6pol can be used to evaluate Pol expression; prot can be used to evaluate protease; p15 for RNAseH; p31 for Integrase; and other appropriate polypeptides for Vif, Vpr, Tat, Rev, Vpu and Nef.

Further, modified polypeptides can also be used, for example, other Env polypeptides include, but are not limited to, for example, native gp160, oligomeric gp140, monomeric gp120 as well as modified and/or synthetic sequences of these polypeptides.

Western Blot analysis can be used to show that cells containing the synthetic expression cassette produce the expected protein, typically at higher per-cell concentrations than cells containing the native expression cassette. The HIV proteins can be seen in both cell lysates and supernatants.

Fractionation of the supernatants from mammalian cells transfected with the synthetic expression cassette can be used to show that the cassettes provide superior production of HIV proteins and relative to the wild-type sequences.

Efficient expression of these HIV-containing polypeptides in mammalian cell lines provides the following benefits: the polypeptides are free of baculovirus contaminants; production by established methods approved by the FDA; increased purity; greater yields (relative to native coding sequences); and a novel method of producing the Sub HIV-containing polypeptides in CHO cells which is not feasible in the absence of the increased expression obtained using the constructs of the present invention. Exemplary Mammalian cell lines include, but are not limited to, BHK, VERO, HT1080, 293, 293T, RD, COS-7, CHO, Jurkat, HUT, SUPT, C8166, MOLT4/clone8, MT-2, MT-4, H9, PM1, CEM, and CEMX174 (such cell lines are available, for example, from the A.T.C.C.).

The desired polypeptide encoding sequences can be cloned into any number of commercially available vectors to generate expression of the polypeptide in an appropriate host system. These systems include, but are not limited to, the following:

baculovirus expression {Reilly, P.R., *et al.*, BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992); Beames, *et al.*, *Biotechniques* 11:378 (1991); Pharmingen; Clontech, Palo Alto, CA)}, vaccinia expression {Earl, P. L., *et al.*, "Expression of proteins in mammalian cells using vaccinia" In *Current Protocols in Molecular Biology* (F. M. Ausubel, *et al.* Eds.), Greene Publishing Associates & Wiley Interscience, New York (1991); Moss, B., *et al.*, U.S. Patent Number 5,135,855, issued 4 August 1992}, expression in bacteria {Ausubel, F.M., *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley and Sons, Inc., Media PA; Clontech}, expression in yeast {Rosenberg, S. and Tekamp-Olson, P., U.S. Patent No. RE35,749, issued, March 17, 1998; Shuster, J.R., U.S. Patent No. 5,629,203, issued May 13, 1997; Gellissen, G., *et al.*, *Antonie Van Leeuwenhoek*, 62(1-2):79-93 (1992); Romanos, M.A., *et al.*, *Yeast* 8(6):423-488 (1992); Goeddel, D.V., *Methods in Enzymology* 185 (1990); Guthrie, C., and G.R. Fink, *Methods in Enzymology* 194 (1991)}, expression in mammalian cells {Clontech; Gibco-BRL, Ground Island, NY; *e.g.*, Chinese hamster ovary (CHO) cell lines (Haynes, J., *et al.*, *Nuc. Acid. Res.* 11:687-706 (1983); 1983, Lau, Y.F., *et al.*, *Mol. Cell. Biol.* 4:1469-1475 (1984); Kaufman, R. J., "Selection and coamplification of heterologous genes in mammalian cells," in *Methods in Enzymology*, vol. 185, pp537-566. Academic Press, Inc., San Diego CA (1991)}, and expression in plant cells {plant cloning vectors, Clontech Laboratories, Inc., Palo Alto, CA, and Pharmacia LKB Biotechnology, Inc., Pistcataway, NJ; Hood, E., *et al.*, *J. Bacteriol.* 168:1291-1301 (1986); Nagel, R., *et al.*, *FEMS Microbiol. Lett.* 67:325 (1990); An, *et al.*, "Binary Vectors", and others in Plant Molecular Biology Manual A3:1-19 (1988); Miki, B.L.A., *et al.*, pp.249-265, and others in Plant DNA Infectious Agents (Hohn, T., *et al.*, eds.) Springer-Verlag, Wien, Austria, (1987); *Plant Molecular Biology: Essential Techniques*, P.G. Jones and J.M. Sutton, New York, J. Wiley, 1997; Miglani, Gurbachan *Dictionary of Plant Genetics and Molecular Biology*, New York, Food Products Press, 1998; Henry, R. J., *Practical Applications of Plant Molecular Biology*, New York, Chapman & Hall, 1997}.

In addition to the mammalian, insect, and yeast vectors, the synthetic expression cassettes of the present invention can be incorporated into a variety of

expression vectors using selected expression control elements. Appropriate vectors and control elements for any given cell can be selected by one having ordinary skill in the art in view of the teachings of the present specification and information known in the art about expression vectors.

For example, a synthetic coding sequence can be inserted into a vector that includes control elements operably linked to the desired coding sequence, which allow for the expression of the coding sequence in a selected cell-type. For example, typical promoters for mammalian cell expression include the SV40 early promoter, a CMV promoter such as the CMV immediate early promoter (a CMV promoter can include intron A), RSV, HIV-Ltr, the mouse mammary tumor virus LTR promoter (MMLV-ltr), the adenovirus major late promoter (Ad MLP), and the herpes simplex virus promoter, among others. Other nonviral promoters, such as a promoter derived from the murine metallothionein gene, will also find use for mammalian expression. Typically, transcription termination and polyadenylation sequences will also be present, located 3' to the translation stop codon. Preferably, a sequence for optimization of initiation of translation, located 5' to the coding sequence, is also present. Examples of transcription terminator/polyadenylation signals include those derived from SV40, as described in Sambrook, et al., *supra*, as well as a bovine growth hormone terminator sequence. Introns, containing splice donor and acceptor sites, may also be designed into the constructs for use with the present invention (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986).

Enhancer elements may also be used herein to increase expression levels of the mammalian constructs. Examples include the SV40 early gene enhancer, as described in Dijkema et al., *EMBO J.* (1985) 4:761, the enhancer/promoter derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus, as described in Gorman et al., *Proc. Natl. Acad. Sci. USA* (1982b) 79:6777 and elements derived from human CMV, as described in Boshart et al., *Cell* (1985) 41:521, such as elements included in the CMV intron A sequence (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986).

Also included in the invention are expression cassettes, comprising coding sequences and expression control elements that allow expression of the coding regions in a suitable host. The control elements generally include a promoter, translation

initiation codon, and translation and transcription termination sequences, and an insertion site for introducing the insert into the vector. Translational control elements useful in expression of the polypeptides of the present invention have been reviewed by M. Kozak (e.g., Kozak, M., *Mamm. Genome* **7(8)**:563-574, 1996; Kozak, M., *Biochimie* **76(9)**:815-821, 1994; Kozak, M., *J Cell Biol* **108(2)**:229-241, 1989; Kozak, M., and Shatkin, A.J., *Methods Enzymol* **60**:360-375, 1979).

Expression in yeast systems has the advantage of commercial production. Recombinant protein production by vaccinia and CHO cell lines have the advantage of being mammalian expression systems. Further, vaccinia virus expression has several advantages including the following: (i) its wide host range; (ii) faithful post-transcriptional modification, processing, folding, transport, secretion, and assembly of recombinant proteins; (iii) high level expression of relatively soluble recombinant proteins; and (iv) a large capacity to accommodate foreign DNA.

The recombinantly expressed polypeptides from immunogenic HIV polypeptide-encoding expression cassettes are typically isolated from lysed cells or culture media. Purification can be carried out by methods known in the art including salt fractionation, ion exchange chromatography, gel filtration, size-exclusion chromatography, size-fractionation, and affinity chromatography. Immunoaffinity chromatography can be employed using antibodies generated based on, for example, HIV antigens. Isolation of oligomeric forms of HIV envelope protein has been previously described (see, e.g., PCT International Application No. WO/00/39302).

Advantages of expressing the proteins of the present invention using mammalian cells include, but are not limited to, the following: well-established protocols for scale-up production; cell lines are suitable to meet good manufacturing process (GMP) standards; culture conditions for mammalian cells are known in the art.

2.3.5 Immunogenicity Enhancing Components for Use with the Polypeptide Component of the Present Invention

Compositions of the present invention for generating an immune response in a mammal, for example, comprising a polynucleotide component and a polypeptide component, can include various excipients, adjuvants, carriers, auxiliary substances,

modulating agents, and the like. The polypeptide component of the compositions of the present invention include an amount of the polypeptide sufficient to mount an immunological response. An appropriate effective amount can be determined by one of skill in the art.

The polypeptide component may comprise a carrier wherein the carrier is a molecule that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycollic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Examples of particulate carriers include those derived from polymethyl methacrylate polymers, as well as microparticles derived from poly(lactides) and poly(lactide-co-glycolides), known as PLG. See, e.g., Jeffery et al., *Pharm. Res.* (1993) 10:362-368; McGee JP, et al., *J Microencapsul.* **14**(2):197-210, 1997; O'Hagan DT, et al., *Vaccine* **11**(2):149-54, 1993. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc., as well as toxins derived from *E. coli*.

Adjuvants may also be used to enhance the effectiveness of the compositions. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) Ribi™ adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene,

0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); (3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) oligonucleotides or polymeric molecules encoding immunostimulatory CpG motifs (Davis, H.L., et al., *J. Immunology* **160**:870-876, 1998; Sato, Y. et al., *Science* **273**:352-354, 1996) or complexes of antigens/oligonucleotides {Polymeric molecules include double and single stranded RNA and DNA, and backbone modifications thereof, for example, methylphosphonate linkages; or (7) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., PCT International Publication Nos. WO/93/13202 and WO/92/19265); (8) Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acteyl-normuramyl-L-alanyl-D-isogluatme (nor-MDP), N-acetylmuramyl-L-alanyl-D-isogluatminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.; (9) Iscomatrix (CSL Limited,Victoria, Australia; also, see, e.g., Morein B, Bengtsson KL, "Immunomodulation by iscoms, immune stimulating complexes," Methods. Sep;19(1):94-102, 1999) and (10) other substances that act as immunostimulating agents to enhance the effectiveness of the composition (e.g., Alum and CpG oligonucleotides).

Preferred adjuvants include, but are not limited to, MF59 and Iscomatrix.

Dosage treatment with the polypeptide component of the immune stimulating compositions of the present invention may be a single dose schedule or a multiple

dose schedule. A multiple dose schedule is one in which a primary course of vaccination may be with 1-10 separate doses, followed by other doses given at subsequent time intervals, chosen to maintain and/or reinforce the immune response, for example at 1-4 months for a second dose, and if needed, a subsequent dose(s) after several months. The dosage regimen will also, at least in part, be determined by the need of the subject and be dependent on the judgment of the practitioner.

Direct delivery of the polypeptide component of the immune-response generating compositions of the present invention is generally accomplished, with or without adjuvants, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). The polypeptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, and needle-less injection. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of polypeptides may also be combined with administration of adjuvants or other substances.

2.3.6 IMMUNOMODULATORY MOLECULES

In some embodiments of the present invention, gene transfer vectors can be constructed to encode a cytokine or other immunomodulatory molecule. For example, nucleic acid sequences encoding native IL-2 and gamma-interferon can be obtained as described in US Patent Nos. 4,738,927 and 5,326,859, respectively, while useful muteins of these proteins can be obtained as described in U.S. Patent No. 4,853,332. Nucleic acid sequences encoding the short and long forms of mCSF can be obtained as described in US Patent Nos. 4,847,201 and 4,879,227, respectively. In particular aspects of the invention, retroviral vectors expressing cytokine or immunomodulatory genes can be produced (e.g., PCT International Publication No. WO/94/02951, entitled "Compositions and Methods for Cancer Immunotherapy").

Examples of suitable immunomodulatory molecules for use herein include the following: IL-1 and IL-2 (Karupiah et al. (1990) *J. Immunology* 144:290-298, Weber

et al. (1987) *J. Exp. Med.* 166:1716-1733, Gansbacher et al. (1990) *J. Exp. Med.* 172:1217-1224, and U.S. Patent No. 4,738,927); IL-3 and IL-4 (Tepper et al. (1989) *Cell* 57:503-512, Golumbek et al. (1991) *Science* 254:713-716, and U.S. Patent No. 5,017,691); IL-5 and IL-6 (Brakenhof et al. (1987) *J. Immunol.* 139:4116-4121, and PCT International Publication No. WO 90/06370); IL-7 (U.S. Patent No. 4,965,195); IL-8, IL-9, IL-10, IL-11, IL-12, and IL-13 (*Cytokine Bulletin*, Summer 1994); IL-14 and IL-15; alpha interferon (Finter et al. (1991) *Drugs* 42:749-765, U.S. Patent Nos. 4,892,743 and 4,966,843, PCT International Publication No. WO 85/02862, Nagata et al. (1980) *Nature* 284:316-320, Familletti et al. (1981) *Methods in Enz.* 78:387-394, Twu et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:2046-2050, and Faktor et al. (1990) *Oncogene* 5:867-872); beta-interferon (Seif et al. (1991) *J. Virol.* 65:664-671); gamma-interferons (Radford et al. (1991) *The American Society of Hepatology* 20082015, Watanabe et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:9456-9460, Gansbacher et al. (1990) *Cancer Research* 50:7820-7825, Maio et al. (1989) *Can. Immunol. Immunother.* 30:34-42, and U.S. Patent Nos. 4,762,791 and 4,727,138); G-CSF (U.S. Patent Nos. 4,999,291 and 4,810,643); GM-CSF (PCT International Publication No. WO 85/04188).

Immunomodulatory factors may also be agonists, antagonists, or ligands for these molecules. For example, soluble forms of receptors can often behave as antagonists for these types of factors, as can mutated forms of the factors themselves.

Nucleic acid molecules that encode the above-described substances, as well as other nucleic acid molecules that are advantageous for use within the present invention, may be readily obtained from a variety of sources, including, for example, depositories such as the American Type Culture Collection, or from commercial sources such as British Bio-Technology Limited (Cowley, Oxford England). Representative examples include BBG 12 (containing the GM-CSF gene coding for the mature protein of 127 amino acids), BBG 6 (which contains sequences encoding gamma interferon), A.T.C.C. Deposit No. 39656 (which contains sequences encoding TNF), A.T.C.C. Deposit No. 20663 (which contains sequences encoding alpha-interferon), A.T.C.C. Deposit Nos. 31902, 31902 and 39517 (which contain sequences encoding beta-interferon), A.T.C.C. Deposit No. 67024 (which contains a sequence

which encodes Interleukin-1b), A.T.C.C. Deposit Nos. 39405, 39452, 39516, 39626 and 39673 (which contain sequences encoding Interleukin-2), A.T.C.C. Deposit Nos. 59399, 59398, and 67326 (which contain sequences encoding Interleukin-3), A.T.C.C. Deposit No. 57592 (which contains sequences encoding Interleukin-4), A.T.C.C. Deposit Nos. 59394 and 59395 (which contain sequences encoding Interleukin-5), and A.T.C.C. Deposit No. 67153 (which contains sequences encoding Interleukin-6).

Plasmids containing cytokine genes or immunomodulatory genes (PCT International Publication Nos. WO 94/02951 and WO 96/21015) can be digested with appropriate restriction enzymes, and DNA fragments containing the particular gene of interest can be inserted into a gene transfer vector using standard molecular biology techniques. (*See, e.g.*, Sambrook et al., *supra.*, or Ausubel et al. (eds) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience).

Polynucleotide sequences coding for the above-described molecules can be obtained using recombinant methods, such as by screening cDNA and genomic libraries from cells expressing the gene, or by deriving the gene from a vector known to include the same. For example, plasmids that contain sequences that encode altered cellular products may be obtained from a depository such as the A.T.C.C., or from commercial sources. Plasmids containing the nucleotide sequences of interest can be digested with appropriate restriction enzymes, and DNA fragments containing the nucleotide sequences can be inserted into a gene transfer vector using standard molecular biology techniques.

Alternatively, cDNA sequences for use with the present invention may be obtained from cells that express or contain the sequences, using standard techniques, such as phenol extraction and PCR of cDNA or genomic DNA. See, e.g., Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA. Briefly, mRNA from a cell which expresses the gene of interest can be reverse transcribed with reverse transcriptase using oligo-dT or random primers. The single stranded cDNA may then be amplified by PCR (see U.S. Patent Nos. 4,683,202, 4,683,195 and 4,800,159, see also *PCR Technology: Principles and Applications for DNA Amplification*, Erlich (ed.), Stockton Press, 1989)) using oligonucleotide primers complementary to sequences on either side of desired sequences.

The nucleotide sequence of interest can also be produced synthetically, rather than cloned, using a DNA synthesizer (*e.g.*, an Applied Biosystems Model 392 DNA Synthesizer, available from ABI, Foster City, California). The nucleotide sequence can be designed with the appropriate codons for the expression product desired. The complete sequence is assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) *Nature* 292:756; Nambair et al. (1984) *Science* 223:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311.

2.4.0 GENERATION OF IMMUNE RESPONSE IN TREATED SUBJECTS

To evaluate efficacy, nucleic acid immunization using the polynucleotide component of the present invention (e.g., two expression cassettes each comprising a coding sequence for gp140, wherein each coding sequence is derived from different HIV subtypes, serotypes, or strains) and antigenic immunization using the polypeptide component of the present invention (e.g., an oligomeric gp140 wherein the coding sequence is derived from one of the HIV subtypes, serotypes, or strains represented in the polynucleotide component) can be performed, for example, as follows.

Example 2 describes methods for the evaluation, in mice, of the immunogenicity of the compositions of the present invention used to induce immune response. The polynucleotide component described comprises two pCMVKM2 each carrying codon optimized coding sequences for gp140 with delV2, the first coding sequence derived from SF162, subtype B, and the second coding sequence derived from TV1, subtype C. The mice are then immunized with oligomeric, codon optimized, gp140 with delV2, derived from SF162, subtype B, polypeptide. Humoral and cellular immune responses are evaluated. The results of these assays are used to show the potency of the polynucleotide/polypeptide immunization methods of the present invention for the generation of an immune response in mice.

Example 3 describes in vivo immunization studies that may be carried out in a variety of laboratory animals (including, mice, guinea pigs, rabbits, rhesus macaques, and baboons). Results of these studies are used to demonstrate the usefulness of the compositions and methods of the invention to generate immune responses, in

particular to generate broad and potent neutralizing activity against diverse HIV strains.

Example 4 describes experiments performed in support of the present invention that evaluated immunogenicity regimens for various HIV polypeptide encoding plasmids used as primes and various HIV polypeptides used as boosts. In the example, the following vectors encoding gp140 proteins were employed: pCMV gp140 dV2 SF162 DNA and pCMV gp140 dV2 TV1 DNA. These vectors comprise expression cassettes that encode gp140 protein derived from two different HIV subtypes, subtype B (SF162) and subtype C (TV1). The V2 loop was deleted in both constructs and the coding sequences were codon optimized for expression in human cells. The specific gp140 polynucleotides have been previously described (e.g., gp140.modSF162.delV2, Figure 6, and gp140.mut7.modSF162.delV2, Figure 7, see also, PCT International Publication No. WO/00/39302; and gp140mod.TV1.delV2, Figure 8, and gp140mod.TV1.mut7.delV2, Figure 9, see also PCT International Publication No. WO/02/04493). The ability of the compositions and methods of the present invention to generate neutralizing antibodies was evaluated. The results of the assays for the presence of neutralizing antibodies are presented in Figure 4 and Figure 5.

Figure 4 summarizes data showing the neutralization titers against HIV-1 SF162 between seven experimental groups. These results demonstrated that all groups showed strong neutralizing activity against the HIV-1 SF162 isolate. Further, neutralizing activity significantly increased at post 4th immunization compared to post 3rd immunizations. For the mixed (B+C) DNA prime and single protein boost, B protein gave a high boost to the mixed gene prime (B+C DNA + B prot), as did the C protein (B+C DNA + C prot). For the mixed DNA prime and protein boost, half dose (50ug) of protein (B+C DNA & prot (1/2)) induced neutralizing activity at least as well as the full dose of 100ug protein (B+C DNA & prot).

Figure 5 summarizes data showing the neutralization titers against HIV-1 TV1 (South African subtype C) between seven experimental groups. These results demonstrated that all groups showed neutralizing activity against HIV1 subtype C TV1 isolate (as expected, because no subtype C DNA or protein was used, the B DNA

+ B prot showed the lowest neutralizing activity). For the mismatched a single DNA prime and a single protein boost (C DNA + B prot), priming with C gene and boosting with B protein showed high titers, as did the B gene and B protein (B DNA + B prot). For the mixed (B+C) DNA prime and single protein boost, use of either B (B+C DNA + B prot) and C (B+C DNA + C prot) proteins had a similar boosting effect.

Comparison of the data presented in Figure 4 and Figure 5 supports the combination methods of the present invention for generating an immune response in a subject, further, for generating neutralizing antibodies in immunized subjects. The data showed that the combination of DNA derived from different subtypes primed broad responses to multiple subtypes. This could be the result of the combined responses to subtype and/or sequence-specific continuous and/or discontinuous immunogenic epitopes as well as responses to the presentation of common conserved eptiopes in the oligomeric V2-deleted Env immunogens employed here. Furthermore, use of a single subtype protein was sufficient to boost broad neutralizing responses when immunity was primed with multiple subtypes of DNA. These results also demonstrated that use of lower doses of proteins mixture can also provide strong immune responses.

Example 5 presents data demonstrating that a subject (in this example chimpanzees) can be immunized with an envelope protein from a first HIV strain of a given subtype (e.g., HIV-1 MN), be boosted with an envelope protein from a second HIV strain of the same subtype (e.g., HIV-1 SF162) and generate neutralizing antibodies against both HIV strains (see, for example Table 11, Example 5). The data in Example 5 supports that the combination methods of the present invention can be used to broadly raise neutralizing antibodies against multiple viral strains from the same subtype. Further, the data presented in Example 4 in combination with the data presented in Example 5 together demonstrate that such HIV strains may be within subtype, or from different subtypes.

These studies demonstrated the usefulness of the compositions (e.g., comprising a polynucleotide component and a polypeptide component) and methods of the invention to generate immune responses, in particular to generate broad and potent neutralizing activity against diverse HIV subtypes and strains. It is readily

apparent that the subject invention can be used to mount an immune response to a wide variety of antigens and hence to treat or prevent infection, particularly HIV infection.

3.0.0 Applications of the Present Invention to HIV

While not desiring to be bound by any particular model, theory, or hypothesis, the following information is presented to provide a more complete understanding of the present invention.

Protection against HIV infection will likely require potent and broadly reactive pre-existing neutralizing antibodies in vaccinated individuals exposed to a virus challenge. Although cellular immune responses are desirable to control viremia in those who get infected, protection against infection has not been demonstrated for vaccine approaches that rely exclusively on the induction of these responses. For this reason, experiments performed in support of the present invention used combination prime-boost approaches that employ a polynucleotide component and a polypeptide component, wherein the polypeptide component encodes, for example, V-deleted envelope antigens from primary HIV isolates (e.g., R5 subtype B (HIV-1_{SF162}) and subtype C (HIV-1_{TVI}) strains), and the polypeptide component comprises at least one of these antigens.

The polynucleotide component of the present invention may be delivered by enhanced DNA or RNA [polylactide co-glycolide (PLG) microparticle formulations or electroporation], adenovirus-based vectors, alphavirus replicons or replicon particles, polynucleotide or particle-based vaccine approaches. Efficient in vivo expression of plasmid encoded genes by electrical permeabilization has been described (see, *e.g.*, Zucchelli et al. (2000) *J. Virol.* 74:11598-11607; Banga et al. (1998) *Trends Biotechnol.* 10:408-412; Heller et al. (1996) *Febs Lett.* 389:225-228; Mathiesen et al. (1999) *Gene Ther.* 4:508-514; Mir et al. (1999) *Proc. Nat'l Acad Sci. USA* 8:4262-4267; Nishi et al. (1996) *Cancer Res.* 5:1050-1055). The polypeptide component of the present invention may be administered, for example, by booster immunizations with Env proteins in MF59 or Iscomatrix adjuvant.

All protein preparations were highly purified and extensively characterized by biophysical and immunochemical methodologies. Results from rabbit immunogenicity studies indicated that broad neutralizing antibody responses could be consistently induced against diverse HIV strains (Example 4). Moreover, using the combination prime-boost vaccine regimens, potent HIV antigen-specific CD4 + and CD8+ T-cell responses may also be generated.

Although any HIV viral protein may also be employed in the practice of the present invention, in a preferred embodiment V1-, V2-, and/or V3-modified/deleted envelope DNA and corresponding polypeptides are good candidates for use in the compositions of the present invention.

One embodiment of this aspect of the present invention may be described generally as follows. Antigens are selected for the vaccine composition(s). Polynucleotides encoding Env polypeptides and Env polypeptides are typically employed in a composition for generating an immunse response comprising a polynucleotide component and a polypeptide component.

A nucleic acid prime is typically followed by at least one polypeptide boost. The boost may, for example, include adjuvanted HIV-derived polypeptides (e.g., analogous to those used for the DNA vaccinations), coding sequences for HIV-derived polypeptides (e.g., analogous to those used for the DNA vaccinations) encoded by a viral vector. Boosts may include further DNA vaccinations, and/or combinations of the foregoing.

Further, different polypeptide antigens may be used in the boost relative to the initial vaccination and visa versa. In addition, the initial nucleic acid vaccination may be a viral vector comprising a DNA expression cassette construct.

Some factors that may be considered in HIV envelope vaccine design are as follows. A fundamental criterion of an effective HIV vaccine is its ability to induce broad and potent neutralizing antibody responses against prevalent HIV strains. The important contribution of neutralizing antibodies in preventing the establishment of HIV, SIV and SHIV infection or delaying the onset of disease is highlighted by several studies. First, the emergence of neutralization-resistant viruses coincides or precedes the development of disease in infected animals (Burns (1993) *J Virol.*

67:4104-13; Cheng-Mayer et al. (1999) *J. Virol.* 73:5294-5300; Narayan et al. (1999) *Virology* 256:54-63). Second, the pre-infusion of high concentrations of potent neutralizing monoclonal antibodies (mAbs) in the blood circulation of macaques, chimpanzees and SCID mice prior to their challenge with HIV, SIV or SHIV viruses, offers protection or delays the onset of disease (Conley et al. (1996) *J. Virol.* 70:6751-6758; Emini et al. (1992) *Nature (London)* 355:728-730; Gauduin et al. (1997) *Nat Med.* 3:1389-93; Mascola et al. (1999) *J Virol.* 73:4009-18; Mascola et al. (2000) *Nature Med.* 6(2):207-210; Baba et al. (2000) *Nature Med.* 6(2):200-206). Similarly, infusion of neutralizing antibodies collected from the serum of HIV-1-infected chimpanzees to naïve pig-tailed macaques protects the latter animals from subsequent viral challenge by SHIV viruses (Shibata et al (1999) *Nature Medicine* 5:204-210). Moreover, envelope-based vaccines have demonstrated protection against infection in non-human primate models. Vaccines that exclude Env-polypeptides generally confer less protective efficacy (see, e.g., Hu, S.L., et al., Recombinant subunit vaccines as an approach to study correlates of protection against primate lentivirus infection, Immunol Lett. Jun;51(1-2):115-9 (1996); Amara, R.R., et al., Critical role for Env as well as Gag-Pol in control of a simian-human immunodeficiency virus 89.6P challenge by a DNA prime/recombinant modified vaccinia virus Ankara vaccine, J Virol. Jun;76(12):6138-46 (2002)).

Monomeric gp120 protein-derived from the SF2 lab strain provided neutralization of HIV-1 lab strains and protection against virus challenges in primate models (Verschoor, E.J., et al., (1999), "Comparison of immunity generated by nucleic acid, MF59 and ISCOM-formulated HIV-1 gp120 vaccines in rhesus macaques," J. Virology 73: 3292-3300). Primary gp120 protein derived from Thai E field strains provided cross-subtype neutralization of lab strains (VanCott, T.C., et al., (1999) "Cross-subtype neutralizing antibodies induced in baboons by a subtype E gp120 immunogen based on an R5 primary human immunodeficiency virus type 1 envelope," J. Virology 73: 4640-4650). Primary sub-type B oligomeric o-gp140 protein provided partial neutralization of subtype B primary (field) isolates (Barnett, S.W., et al. (2001) "The ability of an oligomeric HIV-1 envelope antigen to elicit neutralizing antibodies against primary HIV-1 isolates is improved following the

partial deletion of the second hypervariable region," J. Virology, 75:5526-5540). Primary sub-type B o-gp140 delV2 DNA prime plus protein boost provided potent neutralization of diverse subtype B primary isolates and protection against virus challenge in primate models (Cherpelis, S., et al., (2000) "Vaccine-induced anti-envelope antibodies offer partial protection from SHIV infection to CD8+T-cell depleted rhesus macaques," J. Virology, 75, 1547-1550).

Vaccine strategies for induction of potent, broadly reactive, neutralizing antibodies may be assisted by construction of Envelope polypeptide structures that expose conserved neutralizing epitopes, for example, variable-region modifications/deletions and de-glycosylations, envelope protein-receptor complexes, rational design based on crystal structure (e.g., beta-sheet deletions), and gp41-fusion domain based immunogens.

Stable CHO cell lines for envelope protein production have been developed using optimized envelope polypeptide coding sequences, including, but not limited to, the following: gp120, o-gp140, gp120delV2, o-gp140delV2, gp120delV1V2, o-gp140delV1V2.

Exemplary antigenic compositions and immunogenicity studies in support of the compositions and methods of the present invention are presented in Example 4.

In a first particular aspect of the present invention for compositions for generating an immune response in a mammal, the polynucleotide component of the present invention consists essentially of one polynucleotide encoding an HIV immunogenic polypeptide, and the polypeptide component comprises of one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a different HIV subtype, serotype, or strain than the coding sequence of the immunogenic polypeptide encoded by the polynucleotide component. In this context, the polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide refers to the presence of one polynucleotide encoding one HIV immunogenic polypeptide in the composition. The polynucleotide composition may comprise further components in addition to the HIV immnunogenic polypeptide, such

as immune enhancers, immunoregulatory components, vector sequences (e.g., viral or non-viral), carriers, particles, excipients, expression control sequences, etc. In one embodiment of this aspect of the present invention, the HIV immunogenic polypeptide encoded by the polynucleotide component is derived from a subtype B strain, and at least one coding sequence of an HIV immunogenic polypeptide of the polypeptide component is derived from a subtype C strain.

In one embodiment a composition for generating an immune response in a mammal comprises, a polynucleotide component consisting essentially of a polynucleotide encoding an HIV immunogenic polypeptide derived from a first HIV strain of a first subtype, and a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by the polynucleotide component, provided that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a second HIV strain of the first subtype, wherein the first and second strain are different. In some embodiments of this aspect the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and the polypeptide component does not comprise an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

In a second particular aspect of the present invention for compositions for generating an immune response in a mammal, the polynucleotide component comprises two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV subtypes, serotypes, or strains, and the polypeptide component comprises of one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that (i) if the polypeptide component provides less than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then the HIV immunogenic polypeptides of the polypeptide composition may be derived from the same and/or different HIV subtypes, serotypes, or strains as the HIV immunogenic polypeptides provided by the polynucleotide component, or (ii) if the polypeptide component provides the same or

greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV subtype, serotype, or strain than the HIV immunogenic polypeptides provided by the polynucleotide component.

In one embodiment, the present invention includes a composition for use in generating an immune response in a subject, wherein the composition comprises a polynucleotide encoding an immunogenic HIV polypeptide and an analogous immunogenic HIV polypeptide from a different HIV subtype, serotype, or strain. The polynucleotide encoding an immunogenic HIV polypeptide is used for immunization via delivery of the polynucleotide (e.g., a prime), an analogous immunogenic HIV polypeptide derived from a different HIV subtype, serotype, or strain is used for immunization (e.g., a boost). For example, a DNA molecule is used for nucleic acid immunization, wherein the DNA molecule encodes an HIV envelope polypeptide (i) derived from an HIV subtype C isolate, and (ii) that is codon optimized for expression in mammalian cells. This DNA immunization is followed by a protein boost using an HIV envelope polypeptide derived from an HIV subtype B isolate. Exemplary envelope proteins include, but are not limited to, gp120, gp140, oligomeric gp140, and gp160, including mutated forms thereof (e.g., deletion of the V2 loop). One embodiment of this aspect of the present invention, comprises a composition for generating an immune response in a mammal, the composition comprising: a polynucleotide component having of a first polynucleotide encoding a first HIV immunogenic polypeptide; and a polypeptide component, comprising a second HIV immunogenic polypeptide, wherein the first and second immunogenic HIV polypeptide are derived from different HIV subtypes, serotypes, or strains, and (ii) the first and second immunogenic polypeptides encode analogous HIV polypeptides.

A second embodiment the present invention includes a composition for use in generating an immune response in a subject, wherein the composition comprises a polynucleotide component comprising two or more polynucleotides encoding immunogenic HIV polypeptides, derived from at least two different subtypes, serotypes, or strains, and a polypeptide component having a single, analogous,

immunogenic HIV polypeptides derived from one of the subtypes, serotypes, or strains that is used for boosting. For example, two DNA molecules are used for nucleic acid immunization, wherein the first DNA molecule encodes an HIV envelope polypeptide (i) derived from an HIV subtype C isolate, and (ii) that is codon optimized for expression in mammalian cells, and the second DNA molecule encodes an HIV envelope polypeptide (i) derived from an HIV subtype B isolate, and (ii) that is codon optimized for expression in mammalian cells. This DNA immunization is followed by a protein boost using a single HIV envelope polypeptide (i) derived from an HIV subtype B isolate or an HIV subtype C isolate. Exemplary envelope proteins include, but are not limited to, gp120, gp140, oligomeric gp140, and gp160, including mutated forms thereof (e.g., deletion of the V2 loop). One embodiment of this aspect of the present invention comprises a composition for generating an immune response in a mammal, the composition comprising: a polynucleotide component comprising a first polynucleotide encoding a first immunogenic HIV polypeptide, and a second polynucleotide encoding a second immunogenic HIV polypeptide, wherein (i) the first and second immunogenic HIV polypeptide are derived from different HIV subtypes, and (ii) the first and second immunogenic polypeptides encode analogous HIV polypeptides, and a polypeptide component, having the first HIV immunogenic polypeptide, or the second HIV immunogenic polypeptide, with the proviso that the polypeptide component comprises at least one less HIV immunogenic polypeptide than is encoded by the polynucleotide component.

In another embodiment, a composition for generating an immune response in a mammal comprises a polynucleotide component comprising two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides derived from a first HIV subtype, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains of the first subtype, and a polypeptide component that comprises one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by the polynucleotide component, with the proviso that (i) if the polypeptide component comprises less than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then the HIV immunogenic polypeptides of the

polypeptide composition may be derived from the same and/or different HIV strains of the first subtype, or (ii) if the polypeptide component comprises the same or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain of the first subtype; with the provisos that (i) the polynucleotide component does not encode an HIV immunogenic polypeptide derived from any subtype other than the first subtype, and (ii) the polypeptide component does not comprise an HIV immunogenic polypeptide derived from any subtype other than the first subtype.

The polynucleotide component may comprise further components as described herein (e.g., carriers, vector sequences, control sequences, etc.). The polypeptide component may comprise further components as described herein (e.g., carriers, adjuvants, immunoenhancers, etc.).

In a third aspect, the present invention relates to the use of varied doses of polynucleotides and polypeptides in immunization methods (e.g., prime/boost methods), particularly the methods described herein. Thus, another aspect of the invention provides a method of generating an immune response in a subject comprising administering a polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide derived from a first HIV strain of a first subtype, to a subject under conditions that are compatible with the expression of said polynucleotide in said subject for the production of the encoded HIV immunogenic polypeptide; and, administering a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a second strain of the first subtype, wherein said first HIV strain and said second HIV strain are different. Typically, the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and the polypeptide component does not comprise an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

Another aspect of the present invention provides a method of generating an immune response in a subject comprising administering a polynucleotide component comprising two or more polynucleotides comprising coding sequences for two or more analogous HIV immunogenic polypeptides derived from a first HIV subtype, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains of the first subtype, to a subject under conditions that are compatible with the expression of said polynucleotides in said subject for the production of the encoded HIV immunogenic polypeptides; and, administering a polypeptide component that comprises one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that if the polypeptide component comprises the same number or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain of the first subtype than the HIV immunogenic polypeptides provided by the polynucleotide component. Typically, the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and the polypeptide component does not comprise an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

In a further aspect, the invention provides a method of generating an immune response in a subject comprising

providing a composition comprising

a polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide derived from a first HIVstrain of a first subtype, and

a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a second HIV strain of the first subtype wherein said first and second strains are different;

administering a gene delivery vector comprising the polynucleotide of said polynucleotide component of the composition into said subject under conditions that are compatible with expression of said polynucleotide in said subject for the production of encoded HIV immunogenic polypeptides; and

administering the polypeptide component to said subject.

Yet another aspect of the invention provides a method of generating an immune response in a subject comprising

providing a composition comprising a polynucleotide component comprising two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides derived from a first HIV subtype, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains of the first subtype, and a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that if the polypeptide component comprises the same number or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain of the first subtype than the HIV immunogenic polypeptides provided by the polynucleotide component;

administering one or more gene delivery vectors comprising the polynucleotides of said polynucleotide component of the composition into said subject under conditions that are compatible with expression of said polynucleotides in said subject for the production of encoded HIV immunogenic polypeptides; and

administering the polypeptide component to said subject. Typically, the polynucleotide component of the composition does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and the polypeptide component of the composition does not comprise an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

In any immunization method using, for example, a mixed polynucleotide prime (i.e., two or more polynucleotides encoding immunogenic HIV polypeptides derived from two or more HIV subtypes, serotypes, or strains) in conjunction with a

polypeptide boost the present invention includes using reduced doses of each single component to provide an equivalent immune response to using full doses of each component. In one embodiment, the high threshold of DNA is the maximum tolerable dose of DNA (e.g., about 5 mg to about 10 mg total DNA), the low threshold of DNA is the minimum effective dose (e.g., about 2 ug to about 10 ug total DNA), the high threshold of protein is the maximum tolerable dose of protein (e.g., about 1 mg total protein), the low threshold of protein is the minimum effective dose (e.g., about 2 ug total protein). Experiments performed in support of the present invention demonstrated the efficacy of dividing the total DNA dose among the polynucleotides of the polynucleotide component (e.g., Example 4). Further, experiments performed in support of the present invention (e.g., Example 4) demonstrated the efficacy of dividing the total polypeptide dose among the polypeptides comprising the polypeptide component. The total DNA and total protein are both typically above the low threshold values.

In a preferred embodiment, the total amount of DNA in a given DNA immunization has a high threshold of less than or equal to about 10 mg total DNA and greater than or equal to 1 mg total DNA, and the total amount of protein in a given polypeptide boost has a high threshold of less than or equal to about 200 ug total protein product and greater than or equal to 10 ug of total protein. For example, in an embodiment using a polynucleotide component having two DNA molecules each encoding an immunogenic HIV polypeptide the dose of each DNA molecule per subject may be one milligram of each DNA molecule encoding an immunogenic HIV polypeptide, for a total of 2 mg for the two DNA molecules, or 0.5 mg of each DNA molecule encoding an immunogenic HIV polypeptide, for a total of 1 mg for the two DNA molecules. Dosing with the polypeptide component may be similarly varied, for example, using a polypeptide component having two immunogenic HIV polypeptides the dose of each polypeptide per subject may be 100 micrograms of each immunogenic HIV polypeptide, for a total of 200 ug for the two polypeptides, 50 micrograms of each immunogenic HIV polypeptide, for a total of 100 ug for the two polypeptides, or 25 ug of each immunogenic HIV polypeptide, for a total of 50 ug for

the two polypeptides. As described above, more than two polypeptides may be included in the polypeptide component of the present invention.

In further embodiments, the polynucleotide component of the present invention may comprise one or more gene delivery vectors comprising the polynucleotide(s) encoding immunogenic HIV polypeptide(s). The polypeptide component of the present invention may comprise an adjuvant in addition to the immunogenic polypeptide(s). The present invention also comprises a method for generating an immune response in a subject, the method comprising, administering the polynucleotide composition to the subject under conditions that are compatible with expression of the polynucleotide(s) encoding immunogenic HIV polypeptide(s) in the subject, and administering the polypeptide composition to the subject. The administering of polynucleotide and polypeptide compositions may be concurrent or sequentially. In a preferred embodiment immunization with a polynucleotide component precedes immunization with at least one polypeptide component. Further, a single prime may be followed by multiple boosts or a series of primes and boosts may be used.

Exemplary envelope proteins, and coding sequences thereof, for use in the present invention include, but are not limited to, gp120, gp140, oligomeric gp140, and gp160, including mutated or modified forms thereof (e.g., deletion of the V2 loop, mutations in cleavage sites, or mutations in glycosylation sites). In one embodiment, HIV envelope polypeptides that have been modified to expose the region of their polypeptide that binds to the CCR5 receptor are useful in the practice of the present invention, as well as polynucleotide sequences encoding such polypeptides. From the perspective of humoral immunity, it is useful to generate an immune response that provides neutralization of primary isolates that utilize the CCR5 chemokine co-receptor, which is believed to be important for virus entry (Zhu, T., et al. (1993) Science 261:1179-1181; Fiore, J., et al. (1994) Virology; 204:297-303). These and other exemplary polynucleotide constructs (e.g., a variety of envelope protein coding sequences), methods of making the polynucleotide constructs, corresponding polypeptide products, and methods of making polypeptides useful for HIV immunization have been previously described, for example, in the following: PCT

International Publication Nos.: WO/00/39302; WO/00/39304; WO/02/04493; WO/03/004657; WO/03/004620; and WO/03/020876; US Patent No. 6,602,705; and US Published Patent Application Nos. 20030143248 , and 20020146683.

Although described with reference to HIV subtypes B and C as exemplary subtypes, the compositions and methods of the present invention are applicable to a wide variety of HIV subtypes, serotypes, or strains and immunogenic polypeptides encoded thereby, including but not limited to the following: HIV-1 subtypes, A through K, N and O, the identified CRFs (circulating recombinant forms), and HIV-2 strains and its subtypes. See, e.g., Myers, et al., Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico; Myers, et al., Human Retroviruses and Aids, 1990, Los Alamos, New Mexico: Los Alamos National Laboratory.

Further modifications of Env include, but are not limited to, generating polynucleotides that encode Env polypeptides having mutations and/or deletions therein. For instance, some or all of hypervariable regions, V1, V2, V3, V4 and/or V5 can be deleted or modified as described herein, particularly regions V1, V2, and V3. V1 and V2 regions may mask CCR5 co-receptor binding sites. (See, e.g., Moulard, et al. (2002) Proc. Nat'l Acad. Sci 14:9405-9416). Accordingly, in certain embodiments, some or all of the variable loop regions are deleted, for example to expose potentially conserved neutralizing epitopes. Further, deglycosylation of N-linked sites are also potential targets for modification inasmuch as a high degree of glycosylation also serves to shield potential neutralizing epitopes on the surface of the protein. Additional optional modifications, used alone or in combination with variable region deletes and/or deglycosylation modification, include modifications (e.g., deletions) to the beta-sheet regions (e.g., as described in WO 00/39303), modifications of the leader sequence (e.g., addition of Kozak sequences and/or replacing the modified wild type leader with a native or sequence-modified tpa leader sequence) and/or modifications to protease cleavage sites (e.g., Chakrabarti, et al., (2002) *J. Virol.* 76(11):5357-5368 describing a gp140 Delta CFI containing deletions in the cleavage site, fusogenic domain of gp41, and spacing of heptad repeats 1 and 2 of gp41 that retained native antigenic conformational determinants as defined by binding to known monoclonal antibodies or CD4, oligomer formation, and virus neutralization in vitro).

Various combinations of these modifications can be employed to generate wild-type or synthetic polynucleotide sequences as described herein.

Modification of the Env polypeptide coding sequences may result in (1) improved expression relative to the wild-type coding sequences in a number of mammalian cell lines (as well as other types of cell lines, including, but not limited to, insect cells), and/or (2) improved presentation of neutralizing epitopes. Similar Env polypeptide coding sequences can be obtained, modified and tested for improved expression from a variety of isolates.

Any of the polynucleotides (*e.g.*, expression cassettes) or polypeptides described herein (delivered by any of the methods described above) can also be used in combination with other DNA delivery systems and/or protein delivery systems. Non-limiting examples include co-administration of these molecules, for example, in prime-boost methods where one or more molecules are delivered in a "priming" step and, subsequently, one or more molecules are delivered in a "boosting" step. In certain embodiments, the delivery of one or more nucleic acid-containing compositions is followed by delivery of one or more nucleic acid-containing compositions and one or more polypeptide-containing compositions (*e.g.*, polypeptides comprising HIV antigens). In other embodiments, multiple nucleic acid "primes" (of the same or different nucleic acid molecules) can be followed by multiple polypeptide "boosts" (of the same or different polypeptides). Other examples include multiple nucleic acid administrations and multiple polypeptide administrations.

In any method involving co-administration, the various compositions can be delivered in any order. Thus, in embodiments including delivery of multiple different compositions or molecules, the nucleic acids need not be all delivered before the polypeptides. For example, the priming step may include delivery of one or more polypeptides and the boosting comprises delivery of one or more nucleic acids and/or one more polypeptides. Multiple polypeptide administrations can be followed by multiple nucleic acid administrations or polypeptide and nucleic acid administrations can be performed in any order. Thus, one or more or the nucleic acid molecules (*e.g.*, expression cassettes) described herein and one or more of the polypeptides described herein can be co-administered in any order and via any administration routes.

Therefore, any combination of polynucleotides and polypeptides described herein can be used to elicit an immune reaction.

In addition, following prime-boost regimes (such as those of the present invention described herein) may be beneficial to help reduce viral load in infected subjects, as well as possibly slow or prevent progression of HIV-related disease (relative to untreated subjects).

EXPERIMENTAL

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

Example 1

Generation of Synthetic Expression Cassettes

A. Generating Synthetic Polynucleotides

The polynucleotide sequences used in the practice of the present invention are typically manipulated to maximize expression of their gene products in a desired host or host cell. Following here is some exemplary guidance concerning codon optimization and functional varients of HIV polypeptides. The order of the following steps may vary.

First, the HIV-1 codon usage pattern may be modified so that the resulting nucleic acid coding sequence is comparable to codon usage found in highly expressed human genes. The HIV codon usage reflects a high content of the nucleotides A or T of the codon-triplet. The effect of the HIV-1 codon usage is a high AT content in the DNA sequence that results in a high AU content in the RNA and in a decreased translation ability and instability of the mRNA. In comparison, highly expressed human codons prefer the nucleotides G or C. Wild-type polynucleotide sequences

encoding polypeptides are typically modified to be comparable to codon usage found in highly expressed human genes.

Second, for some genes variants are created (e.g., mutated forms of the wild-type polypeptide). In the following table (Table 2) mutations affecting the activity of several HIV genes are disclosed.

Table 2

Gene	"Region"	Exemplary Mutations
Pol	prot	**Att** = Reduced activity by attenuation of Protease (Thr26Ser) (e.g., Konvalinka et al., 1995, J Virol 69: 7180-86) **Ina** = Mutated Protease, nonfunctional enzyme (Asp25Ala)(e.g., Konvalinka et al., 1995, J Virol 69: 7180-86)
	RT	**YM** = Deletion of catalytic center (YMDD_AP; SEQ ID NO:7) (e.g., Biochemistry, 1995, 34, 5351, Patel et. al.) **WM** = Deletion of primer grip region (WMGY_PI; SEQ ID NO:8) (e.g., J Biol Chem, 272, 17, 11157, Palaniappan, et. al., 1997)
	RNase	no direct mutations, RnaseH is affected by "WM" mutation in RT
	Integrase	1) Mutation of HHCC domain, Cys40Ala (e.g., Wiskerchen et. al., 1995, J Virol, 69: 376). 2.) Inactivation catalytic center, Asp64Ala, Asp116Ala, Glu152Ala (e.g., Wiskerchen et. al., 1995, J Virol, 69: 376). 3) Inactivation of minimal DNA binding domain (MDBD), deletion of Trp235(e.g., Ishikawa et. al., 1999, J Virol, 73: 4475). Constructs int.opt.mut.SF2 and int.opt.mut_C (South Africa TV1) both contain all these mutations (1, 2, and 3)

Gene	"Region"	Exemplary Mutations
Env		Mutations in cleavage site (e.g., Earl et al. (1990) *PNAS USA* 87:648-652; Earl et al. (1991) *J. Virol.* 65:31-41). Mutations in glycosylation site (*e.g.*, GM mutants, for example, change Q residue in V1 and/or V2 to N residue; may also be designated by residue altered in sequence) Deletions or modifications of the V1, V2, V3, V4 or V5 regions or combinations thereof. (See e.g., US 6602705) Deletions or modifications of the β-sheets regions. (See e.g., WO 00/39303)
Tat		Mutants of Tat in transactivation domain (e.g., Caputo et al., 1996, Gene Ther. 3:235), e.g., cys22 mutant (Cys22Gly), cys37 mutant (Cys37Ser), and double mutants
Rev		Mutations in Rev domains (e.g., Thomas et al., 1998, J Virol. 72:2935-44), e.g., mutation in RNA binding-nuclear localization ArgArg38,39AspLeu, mutations in activation domain LeuGlu78,79AspLeu = **M10**
Nef		Mutations of myristoylation signal and in oligomerization domain, for example: 1. Single point mutation myristoylation signal: Gly-to-Ala 2. Deletion of N-terminal first 18 (sub-type B, e.g., SF162) or 19 (sub-type C, e.g., South Africa clones) amino acids. (e.g., Peng and Robert-Guroff, 2001, Immunol Letters 78: 195-200) Single point mutation oligomerization: (e.g., Liu et al., 2000, J Virol 74: 5310-19) Mutations affecting (1) infectivity (replication) of HIV-virions and/or (2) CD4 down regulation. (*e.g.*, Lundquist et al. (2002) *J Virol.* 76(9):4625-33)
Vif		Mutations of Vif: e.g., Simon et al., 1999, J Virol 73:2675-81

Gene	"Region"	Exemplary Mutations
Vpr		Mutations of Vpr: e.g., Singh et al., 2000, J Virol 74: 10650-57
Vpu		Mutations of Vpu: e.g., Tiganos et al., 1998, Virology 251: 96-107

Exemplary polynucleotides comprising some of these mutations have been previously described (see, e.g., PCT International Publication Nos.: WO/00/39302; WO/00/39303; WO/00/39304; WO/02/04493; WO/03/004657; WO/03/004620; and WO/03/020876). Reducing or eliminating the function of the associated gene products can be accomplished employing the teachings set forth in the above table, in view of the teachings of the present specification.

In one aspect, the present invention comprises *Env* coding sequences that include, but are not limited to, polynucleotide sequences encoding the following HIV-encoded polypeptides: gp160, gp140, and gp120 (see, e.g., U.S. Patent No. 5,792,459 for a description of the HIV-1_{SF2} ("SF2") Env polypeptide). The relationships between these polypeptides is shown schematically in Figure 3 (in the figure: the polypeptides are indicated as lines, the amino and carboxy termini are indicated on the gp160 line; the open circle represents the oligomerization domain; the open square represents a transmembrane spanning domain (TM); and "c" represents the location of a cleavage site, in gp140.mut the "X" indicates that the cleavage site has been mutated such that it no longer functions as a cleavage site). The polypeptide gp160 includes the coding sequences for gp120 and gp41. The polypeptide gp41 is comprised of several domains including an oligomerization domain (OD) and a transmembrane spanning domain (TM). In the native envelope, the oligomerization domain is required for the non-covalent association of three gp41 polypeptides to form a trimeric structure: through non-covalent interactions with the gp41 trimer (and itself), the gp120 polypeptides are also organized in a trimeric structure. A cleavage site (or cleavage sites) exists approximately between the polypeptide sequences for gp120 and the polypeptide sequences corresponding to gp41. This cleavage site(s) can be mutated to prevent cleavage at the site. The resulting gp140 polypeptide corresponds to a

truncated form of gp160 where the transmembrane spanning domain of gp41 has been deleted. This gp140 polypeptide can exist in both monomeric and oligomeric (*i.e.* trimeric) forms by virtue of the presence of the oligomerization domain in the gp41 moiety. In the situation where the cleavage site has been mutated to prevent cleavage and the transmembrane portion of gp41 has been deleted the resulting polypeptide product is designated "mutated" gp140 (e.g., gp140.mut). As will be apparent to those in the field, the cleavage site can be mutated in a variety of ways. (See, also, e.g., PCT International Publication Nos. WO 00/39302 and WO/02/04493).

Wild-type HIV coding sequences (*e.g.*, Gag, Env, Pol, tat, rev, nef, vpr, vpu, vif, etc.) can be selected from any known HIV isolate and these sequences manipulated to maximize expression of their gene products following the teachings of the present invention. The wild-type coding region maybe modified in one or more of the following ways: sequences encoding hypervariable regions of Env, particularly V1 and/or V2 are deleted, and/or mutations are introduced into sequences, for example, encoding the cleavage site in Env to abrogate the enzymatic cleavage of oligomeric gp140 into gp120 monomers. (See, e.g., Earl et al. (1990) *PNAS USA* 87:648-652; Earl et al. (1991) *J. Virol.* 65:31-41). In yet other embodiments, hypervariable region(s) are deleted, N-glycosylation sites are removed and/or cleavage sites mutated. As discussed above, different mutations may be introduced into the coding sequences of different genes (see, e.g., Table 2).

To create the synthetic coding sequences of the present invention the gene cassettes are designed to comprise the entire coding sequence of interest. Synthetic gene cassettes are constructed by oligonucleotide synthesis and PCR amplification to generate gene fragments. Primers are chosen to provide convenient restriction sites for subcloning. The resulting fragments are then ligated to create the entire desired sequence which is then cloned into an appropriate vector. The final synthetic sequences are (i) screened by restriction endonuclease digestion and analysis,(ii) subjected to DNA sequencing in order to confirm that the desired sequence has been obtained and (iii) the identity and integrity of the expressed protein confirmed by SDS-PAGE and Western blotting. The synthetic coding sequences are assembled at

Chiron Corp. (Emeryville, CA) or by the Midland Certified Reagent Company (Midland, Texas).

Percent identity to the synthetic sequences of the present invention can be determined, for example, using the Smith-Waterman search algorithm (Time Logic, Incline Village, NV), with the following exemplary parameters: weight matrix = nuc4x4hb; gap opening penalty = 20, gap extension penalty = 5, reporting threshold = 1; alignment threshold = 20.

Various forms of the different embodiments of the present invention (*e.g.*, constructs) may be combined.

Some exemplary embodiments of synthetic polynucleotides useful in the practice of the present invention are discussed in Example 4 and presented in Figure 6 to Figure 19.

B. Creating Expression Cassettes Comprising the Synthetic Polynucleotides of the Present Invention

The synthetic DNA fragments of the present invention may be cloned into a number of viral or non-viral expression vectors. For example, polynucleotides used in the practice of the present invention may be cloned into the following non-viral expression vectors: (i) pCMVKm2, for transient expression assays and DNA immunization studies, the pCMVKm2 vector was derived from pCMV6a (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986) and comprises a kanamycin selectable marker, a ColE1 origin of replication, a CMV promoter enhancer and Intron A, followed by an insertion site for the synthetic sequences described below followed by a polyadenylation signal derived from bovine growth hormone -- the pCMVKm2 vector differs from the pCMV-link vector only in that a polylinker site was inserted into pCMVKm2 to generate pCMV-link; (ii) pESN2dhfr and pCMVPLEdhfr (also known as pCMVIII), for expression in Chinese Hamster Ovary (CHO) cells; and, (iii) pAcC13, a shuttle vector for use in the Baculovirus expression system (pAcC13, was derived from pAcC12 which was described by Munemitsu S., et al., *Mol Cell Biol.* 10(11):5977-5982, 1990). See, also PCT International Publication Nos. WO

00/39302, WO 00/39303, WO 00/39304, WO 02/04493 for a description of these vectors.

Briefly, construction of pCMVPLEdhfr (pCMVIII) was as follows. To construct a DHFR cassette, the EMCV IRES (internal ribosome entry site) leader was PCR-amplified from pCite-4a+ (Novagen, Inc., Milwaukee, WI) and inserted into pET-23d (Novagen, Inc., Milwaukee, WI) as an *Xba-Nco* fragment to give pET-EMCV. The *dhfr* gene was PCR-amplified from pESN2dhfr to give a product with a Gly-Gly-Gly-Ser spacer in place of the translation stop codon and inserted as an *Nco-Bam*H1 fragment to give pET-E-DHFR. Next, the attenuated *neo* gene was PCR amplified from a pSV2Neo (Clontech, Palo Alto, CA) derivative and inserted into the unique *Bam*H1 site of pET-E-DHFR to give pET-E-DHFR/$Neo_{(m2)}$. Then, the bovine growth hormone terminator from pCDNA3 (Invitrogen, Inc., Carlsbad, CA) was inserted downstream of the *neo* gene to give pET-E-DHFR/$Neo_{(m2)}$BGHt. The EMCV-*dhfr/neo* selectable marker cassette fragment was prepared by cleavage of pET-E-DHFR/$Neo_{(m2)}$BGHt. The CMV enhancer/promoter plus Intron A was transferred from pCMV6a (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986) as a *Hin*dIII-*Sal*1 fragment into pUC19 (New England Biolabs, Inc., Beverly, MA). The vector backbone of pUC19 was deleted from the Nde1 to the Sap1 sites. The above described DHFR cassette was added to the construct such that the EMCV IRES followed the CMV promoter to produce the final construct. The vector also contained an amp^r gene and an SV40 origin of replication.

Expression vectors of the present invention may comprise one or more polynucleotide sequence encoding immunogenic polypeptides. When the expression cassette contains more than one coding sequence the coding sequences may all be in-frame to generate one polyprotein; alternatively, the more than one polypeptide coding sequences may comprise a polycistronic message where, for example, an IRES is placed 5' to each polypeptide coding sequence; further, multiple promoters may be present to direct the expression of multiple coding sequences.

Example 2

In Vivo Immunogenicity in Mice of Synthetic HIV Expression Cassettes and Polypeptides Encoded Thereby

A. Immunization

To evaluate the immunogenicity of the compositions of the present invention used to induce immune response, a mouse study may be performed. The polynucleotide component (e.g., two pCMVlink-based plasmids each carrying codon optimized coding sequences for gp140 with delV2, the first coding sequence derived from SF162, subtype B, and the second coding sequence derived from TV1, subtype C), is diluted in a total injection volume of 100 µl using varying doses of DNA (0.02 – 200µg). To overcome possible negative dilution effects of the diluted DNA, the total DNA concentration in each sample can be adjusted using the vector (e.g., pCMVlink) alone. Groups of 10-12 Balb/c mice (Charles River, Boston, MA) are intramuscularly immunized (50 µl per leg, intramuscular injection into the *tibialis anterior*) using varying dosages.

The mice are then immunized with oligomeric, codon optimized, gp140 with delV2, derived from SF162, subtype B, polypeptide at intervals following the DNA immunization using appropriate concentrations of polypeptide.

B. Humoral Immune Response

The humoral immune response is checked with a suitable anti-HIV antibody ELISAs (enzyme-linked immunosorbent assays) of the mice sera 0 and 2-4 week intervals post immunization.

The antibody titers of the sera are determined by anti-HIV antibody ELISA. Briefly, sera from immunized mice are screened for antibodies directed against HIV envelope protein. ELISA microtiter plates are coated with 0.2 µg of HIV envelope gp140 protein per well overnight and washed four times; subsequently, blocking is done with PBS-0.2% Tween (Sigma) for 2 hours. After removal of the blocking solution, 100 µl of diluted mouse serum is added. Sera are tested at 1/25 dilutions and by serial 3-fold dilutions, thereafter. Microtiter plates are washed four times and

incubated with a secondary, peroxidase-coupled anti-mouse IgG antibody (Pierce, Rockford, IL). ELISA plates are washed and 100 µl of 3, 3', 5, 5'-tetramethyl benzidine (TMB; Pierce) is added per well. The optical density of each well is measured after 15 minutes. The titers reported are the reciprocal of the dilution of serum that gave a half-maximum optical density (O.D.).

The results of these assays are used to show the potency of the polynucleotide/polypeptide immunization methods of the present invention for the generation of an immune response in mice.

C. Cellular Immune Response

The frequency of specific cytotoxic T-lymphocytes (CTL) is evaluated by a standard chromium release assay of peptide pulsed Balb/c mouse CD4 cells. HIV protein-expressing vaccinia virus infected CD-8 cells may be used as a positive control (vv-protein). Briefly, spleen cells (Effector cells, E) are obtained from the BALB/c mice (immunized as described above). The cells are cultured, restimulated, and assayed for CTL activity against, e.g., Envelope peptide-pulsed target cells (see, e.g., Doe, B., and Walker, C.M., *AIDS* 10(7):793-794, 1996, for a general description of the assay). Cytotoxic activity is measured in a standard ^{51}Cr release assay. Target (T) cells are cultured with effector (E) cells at various E:T ratios for 4 hours and the average cpm from duplicate wells is used to calculate percent specific ^{51}Cr release. Antigen specific T cell responses in immunized mice can also be measured by flow cytometric determinations of intracellular cytokine production (Cytokine flow Cytometry or "CFC") as described by zur Megede, J., et al.,, in Expression and immunogenicity of sequence-modified human immunodeficiency virus type 1 subtype B pol and gagpol DNA vaccines, J Virol. 77:6197-207 (2003).

Cytotoxic T-cell (CTL) or CFC activity is measured in splenocytes recovered from the mice immunized with HIV DNA constructs and polypeptides as described herein. Effector cells from the immunized animals typically exhibit specific lysis of HIV peptide-pulsed SV-BALB (MHC matched) targets cells indicative of a CTL response. Target cells that are peptide-pulsed and derived from an MHC-unmatched mouse strain (MC57) are not lysed. The results of the CTL or CFC assays are used to

show the potency of the polynucleotide/polypeptide immunization methods of the present invention for induction of cytotoxic T-lymphocyte (CTL) responses by DNA immunization.

Example 3

In Vivo Immunogenicity Studies

A. General Immunization Methods

To evaluate the immune response generated using the compositions (comprising a polynucleotide component and a polypeptide component) and methods of the present invention, studies using guinea pigs, rabbits, mice, rhesus macaques and/or baboons may be performed. The studies are typically structured as shown in the following table (Table 3) and can be carried out using, for example, the following components: subtype B DNA -- pCMVlink carrying a codon optimized coding sequences for gp140 with delV2, the coding sequence derived from SF162, subtype B; subtype C DNA -- pCMVlink carrying a codon optimized coding sequences for gp140 with delV2, the coding sequence derived from TV1, subtype C; subtype B protein -- oligomeric, codon optimized, gp140 with delV2, derived from SF162, subtype B polypeptide; and subtype C protein -- oligomeric, codon optimized, gp140 with delV2, derived from TV1, subtype C polypeptide.

Table 3

DNA Immunization	Protein Immunization			
	Subtype B	Subtype C	Subtype B & C (1X)	Subtype B & C (2X)
Subtype B	X	X	X	X
Subtype C	X	X	X	X
Subtype B & C (1X)	X	X	X	X
Subtype B & C (2X)	X	X	X	X

The immunizations may use single or multiple DNA immunizations and single or multiple protein immunizations. The immunizations in the above table exemplify

the following methods: prime/boost regimens (Subtype B DNA/Subtype B protein; Subtype C DNA/Subtype C protein); mixed prime/boost, single DNA prime and single -protein boost (Subtype B DNA/Subtype C protein; Subtype C DNA/Subtype B protein); mixed DNA prime single protein boost (Subtype B & C DNA/Subtype B protein; Subtype B & C DNA/Subtype C protein); single DNA prime mixed protein boost (Subtype B DNA/Subtype B & C protein; Subtype C DNA/Subtype B & C protein); and mixed DNA prime mixed protein boost (Subtype B& C DNA/Subtype B & C protein. The amount of each DNA and /or protein in the mixed samples (i.e, B & C, in this example) can be added at an amount equal to that delivered in the single imuunizations (such that 2X the amount of total DNA and/or protein is delivered) or the amount of each DNA and/or protein in the mixed samples can be adjusted so that the same total amount (1X) of DNA and/or protein is delivered in the mixed and single samples.

In addition to examples in Table 3 exemplifying combinations of polynucleotide component and polypeptide component, other combinations exemplifying two polynucleotide or two polypeptide components can be mentioned. For example, continuing the above example using combinations of HIV subtype B and subtype C immunogens, the present invention also includes single DNA prime and single DNA boost (Subtype B DNA/Subtype C DNA); single protein prime and single protein boost (Subtype B protein/Subtype C protein).

B. Mice

Experiments may be performed in mice following the immunization protocol illustrated in Table 3 and using the methods essentially as described in Example 2.

C. Guinea Pigs

Experiments may be performed in guinea pigs as follows. Groups comprising six guinea pigs each are immunized parenterally (e.g., intramuscularly or intradermally) or mucosally at 0, 4, and 12 weeks with plasmid DNAs comprising expression cassettes comprising one or more HIV immunogenic polypeptide (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. A subset

of the animals are subsequently boosted at approximately 12-24 weeks with a single dose (intramuscular, intradermally or mucosally) of the HIV protein(s) (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. Animals may be boosted subsequently multiple times at 8-16 week intervals with the HIV protein. Antibody titers (geometric mean titers) are measured at two weeks following the third DNA immunization and at two weeks after the protein boost. Results of these studies are used to demonstrate the usefulness of the compositions and methods of the invention to generate immune responses, in particular to generate broad and potent neutralizing activity against diverse HIV strains.

D. Rabbits

Experiments may be performed in rabbits as follows. Rabbits are immunized intramuscularly or intradermally at multiple sites (using needle injection with or without subsequent electroporation, or using a Bioject needless syringe) or mucosally with with plasmid DNAs comprising expression cassettes comprising one or more HIV immunogenic polypeptide (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. A subset of the animals are subsequently boosted with a single dose (intramuscular, intradermally or mucosally) of the HIV protein(s) (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. Animals may be boosted multiple times with the HIV protein. Typically, the compositions of the present invention used to generate immune responses are highly immunogenic and generate substantial antigen binding antibody responses after only 2 immunizations in rabbits. Results of these studies are used to demonstrate the usefulness of the compositions and methods of the invention to generate immune responses, in particular to generate broad and potent neutralizing activity against diverse HIV strains.

E. Rhesus Macaques

Experiments may be performed in rhesus macaques as follows. Rhesus macaques are immunized at approximately 0, 4, 8, and 24 weeks parenterally or mucosally with plasmid DNAs comprising expression cassettes comprising one or

more HIV immunogenic polypeptide (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. Enhanced DNA delivery systems such as use of DNA complexed to PLG microparticles or saline injection of DNA followed by electropoartion can be employed to increase immune response during the DNA priming phase of the immunization regimen. A subset of the animals are subsequently boosted with a single dose (intramuscular, intradermally or mucosally) of the HIV protein(s) (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. Animals may be boosted multiple times generally at 3-6 month intervals with the HIV protein. Typically, the macaques have detectable HIV-specific T-cell responses as measured by CTL assays or Cytokine Flow Cytometry after two or three 1 mg doses of the polynucleotide component. Neutralizing antibodies may also detected. Results of these studies are used to demonstrate the usefulness of the compositions and methods of the invention to generate immune responses, in particular to generate broad and potent neutralizing activity against diverse HIV strains.

F. Baboons

Baboons are immunized 4 times (at approximately weeks 0, 4, 8, and 24) intramuscular, or intradermally, or mucosally with plasmid DNAs comprising expression cassettes comprising one or more HIV immunogenic polypeptide (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. The DNAs can be delivered in saline with or without electroporation, or on PLG microparticles. A subset of the animals are subsequently boosted with a single dose (intramuscular, intradermally or mucosally) of the HIV protein(s) (for example, gp 140 DNAs as descibed in Example 2) as illustrated in Table 3. Animals may be boosted multiple times generally at 3-6 month intervals with the HIV protein.

The animals are bled two-four weeks after each immunization and an HIV antibody ELISA is performed with isolated plasma. The ELISA is performed essentially as described below in Section G except the second antibody-conjugate is typically an anti-human IgG, g-chain specific, peroxidase conjugate (Sigma Chemical Co., St. Louis, MD 63178) used at a dilution of 1:500. Fifty μg/ml yeast extract may

be added to the dilutions of plasma samples and antibody conjugate to reduce non-specific background due to preexisting yeast antibodies in the baboons. Lymphoproliferative responses to are typically observed in baboons post-boosting with HIV-polypeptide Such proliferation results are indicative of induction of T-helper cell functions. Results of these studies are used to demonstrate the usefulness of the compositions and methods of the invention to generate immune responses, in particular to generate broad and potent neutralizing activity against diverse HIV strains.

G. Humoral Immune Response

In any immunized animal model (including the above, as well as, for example, chimpanzees), the humoral immune response is checked in serum specimens from the immunized animals with an anti-HIV antibody ELISAs (enzyme-linked immunosorbent assays) at various times post-immunization. The antibody titers of the sera are determined by anti-HIV antibody ELISA as described above. Briefly, sera from immunized animals are screened for antibodies directed against the HIV polypeptide/protein(s) encoded by the DNA and/or polypeptide used to immunize the animals (e.g., oligomeric gp140). Typically independent ELISA assays are carried out using polypeptides corresponding to each of the subtypes used in the immunization study.

Wells of ELISA microtiter plates are coated overnight with the selected HIV polypeptide/protein and washed four times; subsequently, blocking is done with PBS-0.2% Tween (Sigma) for 2 hours. After removal of the blocking solution, 100 µl of diluted mouse serum is added. Sera are tested at 1/25 dilutions and by serial 3-fold dilutions, thereafter. Microtiter plates are washed four times and incubated with a secondary, peroxidase-coupled anti-mouse IgG antibody (Pierce, Rockford, IL). ELISA plates are washed and 100 µl of 3, 3', 5, 5'-tetramethyl benzidine (TMB; Pierce) was added per well. The optical density of each well is measured after 15 minutes. Titers are typically reported as the reciprocal of the dilution of serum that gave a half-maximum optical density (O.D.).

Cellular immune responses may also be evaluated.

The presence of neutralizing antibodies in the sera is determined essentially as follows: Virus neutralization is measured in 5.25.EGFP.Luc.M7 (M7-luc) cells obtained from Dr. Nathaniel Landau (Salk Institute, San Diego, CA). The format of this assay is essentially the same as the MT-2 assay as described elsewhere (Montefiori et al. (1988) *J. Clin Microbiol.* 26:231-235) except that virus infection is quantified by luciferase reporter gene expression using a commercial luciferase kit (Promega). All serum samples are heat-inactivated for 1 hour at 56°C prior to assay. The virus stocks of the HIV-1 isolates are typically generated in PBMC.

Example 4

Evaluation Of Immunogenicity Regimens For Various HIV Polypeptide Encoding Plasmids Used As Primes And Various HIV Polypeptides Used As Boosts

To evaluate the combination effects of subtype C (TV1) and subtype B (SF162) pg140dV2 DNAs and proteins for DNA prime/boost the following experiments were carried out in rabbits. DNA was gp140mod.TV1.dV2 and gp140mod.SF162.dV2, delivered separately in two plasmids (sources of DNA are described further herein below). Protein was oligomer o-gp140.dV2.TV1 and o-gp140.dV2.SF162 (sources of the proteins are described further herein below). DNA constructs were used for immunization in three doses at schedules of 0, 4, 12 weeks. Proteins were boosted at 12, 24, and 41 weeks. Each rabbit was injected 1.0 ml DNA mixture at two sides IM/Quadriceps, followed by an electroporation procedure (G. Widera, Increased DNA vaccine delivery and immunogenicity by electroporation *in vivo*, J. Immunology, 164, 4635-4640 (2000)). MF59 adjuvanted protein was injected two sites, IM/Glut for 1ml per animal.

All of the genes were sequence-modified to enhance expression of the encoded Env glycoproteins in a Rev-independent fashion and they were subsequently cloned into pCMV-based plasmid vectors for DNA vaccine and protein production applications as described above. The sequences were codon optimized as described herein. Briefly, all the modified envelope genes were cloned into the Chiron pCMVlink plasmid vector, preferably into EcoRI/XhoI sites.

To obtain gp140 polypeptides each of the gp140 contructs (i.e., gp140mod.TV1.mut7.delV2 and gp140.mut7.modSF162.delV2) were used in the following method.

Chinese hamster ovary (CHO) cells were transfected with plasmid DNA encoding the gp140 proteins (e.g., pCMV vector backbone) using Mirus TransIT-LT1 polyamine transfection reagent (Mirus Corporation, Madison WI) according to the manufacturer's instructions and incubated for 96 hours. After 96 hours, media was changed to selective media (F12 special with 250 μg/ml G418) and cells were split 1:5 and incubated for an additional 48 hours. Media was changed every 5-7 days until colonies started forming at which time the colonies were picked, plated into 96 well plates and screened by gp120 Capture ELISA. Positive clones were expanded in 24 well plates and screened several times for Env protein production by Capture ELISA, as described above. After reaching confluency in 24 well plates, positive clones were expanded to T25 flasks (Corning, Corning, NY). These were screened several times after confluency and positive clones were expanded to T75 flasks.

Positive T75 clones were frozen in liquid nitrogen and the highest expressing clones amplified with 0-5 μM methotrexate (MTX)at several concentrations and plated in 100 mm culture dishes. Plates were screened for colony formation and all positive closed were again expanded as described above. Clones were expanded, amplified and screened at each step by gp120 capture ELISA. Positive clones were frozen at each methotrexate level. Highest producing clones were grown in perfusion bioreactors (3L, 100L) for expansion and adaptation to low serum suspension culture conditions for scale-up to larger bioreactors.

The stably transfected CHO cell lines, which express the Env polypeptides, were used to produce gp140 proteins. The proteins were purified, briefly, by using a three-step strategy as previously described (Srivastava, et al., Purification and characterization of oligomeric envelope glycoprotein from a primary r5 subtype B human immunodeficiency virus. J Virol **76**:2835-47 (2002)). First, concentrated cell supernatants were passed over a Galanthus Nivalis-agarose column (GNA; Vector Laboratories, Burlingame, CA). The gp140SF162ΔV2 protein bound to the column, and most contaminating proteins flowed through. The bound protein was eluted with

500 mM methyl mannose pyranoside (MMP). Next, the captured protein was passed over DEAE and CHAP columns.

These methods are applicable to other HIV genes and proteins derived from other HIV subtypes. Further, although this analysis was carried out in rabbits similar analysis may be carried out with other type of animals, for example, as described in Example 3. The immunization weeks can be varied.

The following table (Table 4) lists exemplified procedures used in a comparison of the immunogenicity of subtype B and C polynucleotides encoding envelope polypeptides (in a pCMVlink vector) in various combinations with subtype B and C envelope polypeptides, both individually and as a mixed-subtype vaccine, using electroporation, in rabbits. It will be apparent to one skilled in the art in view of the teachings of the present specification that such methods are equally applicable to any other polynucleotides encoding immunogenic HIV polypeptides and immunogenic HIV polypeptides.

Table 4

Group	Animal #	Imm'n #	Adjuvant	Immunogen	Total Dose	Vol/ Site	Sites/ Animal	Route
1	1-4	1, 2, 3, 4	MF59C	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)
2	5-8	1, 2, 3, 4	Iscomatrix	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)
3	9-12	1, 2, 3	-	pCMV 140 dV2 SF162 DNA	1.0mg	0.50ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)
4	13 -16	1, 2, 3	-	pCMV 140 dV2 SF162 DNA	1.0mg	0.5ml	2	IM/Quad (Needle)
		3, 4	Iscomatrix	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)
5	17-20	1, 2,3,4	MF59C	o-gp140 dV2 TV1	50ug	500ul	2	IM/Glut

Group	Animal #	Imm'n #	Adjuvant	Immunogen	Total Dose	V l/ Site	Sites/ Animal	Route
								(Needle)
6	21-24	1, 2, 3	-	pCMV 140 dV2 TV1 DNA	1.0mg	0.5ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)
7	25-28	1, 2, 3	-	pCMV 140 dV2 SF162 DNA pCMV 140 dV2 TV1 DNA	2.0mg (1.0mg ea.)	0.5ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)
8	29-32	1, 2, 3	-	pCMV 140 dV2 SF162 DNA pCMV 140 dV2 TV1 DNA	2.0mg	0.5ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 TV1	50ug	500ul	2	IM/Glut (Needle)
9	33-36	1, 2, 3	-	pCMV 140 dV2 SF162 DNA pCMV 140 dV2 TV1 DNA	2.0mg	0.50ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 SF162 o-gp140 dV2 TV1	100ug	500ul	2	IM/Glut (Needle)
10	37-40	1, 2, 3	-	pCMV 140 dV2 SF162 DNA pCMV 140 dV2 TV1 DNA	2.0mg	0.5ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 SF162 o-gp140 dV2 TV1	50ug	500ul	2	IM/Glut (Needle)
11	41-44	1, 2,3	-	pCMV 140 dV2 SF162 DNA pCMV 140 dV2 TV1 DNA	1.0mg	0.50ml	2	IM/Quad (Needle)
		3, 4	MF59C	o-gp140 dV2 SF162	50ug	500ul	2	IM/Glut (Needle)

The MF59C adjuvant is a microfluidized emulsion containing 5% squalene, 0.5% Tween 80, 0.5% Span 85, in 10mM citrate pH 6, stored in 10 ml aliquots at 4°C.

The Iscomatrix adjuvant is a quil saporin based adjuvant used for protein delivery (available from, e.g., CSL Limited,Victoria, Australia).

The polynucleotides and polypeptides listed in Table 4 were prepared as described in Table 5.

Table 5

Polynucleotide Construct / Polypeptide	Description
pCMV 140 dV2 SF162 DNA	The plasmid (pCMVlink) contained a synthetic, codon optimized HIV-1 gp140 envelope gene from subtype B strain SF162 (see, gp140.modSF162.delV2, Figure 6, see also PCT International Publication No. WO/00/39302). The gp140 gene comprised the gp120 and gp41 ectodomain. The constructs also contained a deletion in the variable region V2 (dV2).The plasmid construct contained the human CMV enhancer/promoter and Kanamycin resistance gene. Plasmids were prepared by alkaline lysis method and Qiagen purification from DH5--□nE.coli bacteria. Plasmids were stored at –80C until use.
pCMV 140 dV2 TV1 DNA	The plasmid (pCMVlink) contained a synthetic, codon optimized HIV-1 gp140 envelope gene derived from HIV-1 subtype C strain TV1 (see, gp140mod.TV1.delV2, Figure 8, see also PCT International Publication No. WO/02/04493). The structure of the envelope gene and the plasmid was as described above.
o-gp140 dV2 SF162 protein	The subtype B oligomer protein contained five amino acid mutations in the cleavage site in addition to the deletion of V2 region (see, gp140.mut7.modSF162.delV2, Figure 7, see also PCT International Publication No. WO/00/3930). Protein was expressed in CHO cells and purified from the CHO cells. Expression and purification of o-gp140 proteins was described, for example, in PCT International Publication No. WO/00/39302 and Srivastava, et al., J Virol 76:2835-47 (2002).
o-gp140 dV2 TV1 protein	The subtype C oligomer protein contained five amino acid mutations in the cleavage site in

Polynucleotide Construct / Polypeptide	Description
	addition to the deletion of V2 region (see, gp140mod.TV1.mut7.delV2, Figure 9, see also PCT International Publication No. WO/02/04493). Protein was expressed in CHO cells and purified from the CHO cells. Expression and purification of o-gp140 proteins was described, for example, in PCT International Publication No. WO/00/39302 and Srivastava, et al., J Virol 76:2835-47 (2002).

Immunogens were prepared as described in the following table (Table 6) for administration to animals in the various groups.

Table 6

Group	Preparation
1, 5	**Immunization 1-4: Protein Immunization + MF59** Protein doses were 50ug protein per animal. The initial protein was diluted to 0.100 mg/ml in citrate buffer. Stored at –80°C until use. Thawed at room temperature; material was clear with no particulate matter. Added equal volume of MF59C adjuvant to thawed protein and mixed well by inverting the tube. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal. Used material within 1 hour of the addition of adjuvant. Needles were used for injections.
2	**Immunization 1-4: Protein Immunization + Iscomatrix** The stock concentration was 1mg/ml. Immediately before immunizations, 250ul of 1mg/ml Iscomatrix was diluted to 2.5ml of 0.1mg/ml with PBS (CFU U21). Added equal volume (2.5ml) of 0.1mg/ml Iscomatrix into 2.5ml of 0.1mg/ml protein and mixed well. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal.
3- 4, 6	**Immunization 1--3: Subtype B/C plasmid DNA in Saline** The immunogen was provided at 1.0 mg/ml total DNA in sterile 0.9% saline. Stored at –80°C until use. Thawed DNA at room temperature; the material was clear or slightly opaque, with no particulate matter. Immunized each rabbit with 0.5ml DNA mixture per side (IM/Quadriceps), total 2 sides with 1.0ml per animal. Animals were shaved prior to immunization, under sedation of 1x dose IP (by animal weight) of Ketamine-Xylazine (80mg/ml - 4mg/ml). DNA injection used needle. Following the DNA injection, electroporation was administrated

Group	Preparation
	using a 6-needle circular array with 1cm diameter, 1cm needle length. Electroporation pulses were given at 20V/mm, 50ms pulse length, 1 pulse/s.
3, 6	**Immunization 3-4: Protein Immunization** Protein doses were 50ug each SF162 protein per animal. The initial SF162 Protein was diluted to 0.100 mg/ml in citrate buffer. Stored at –80°C until use. Thawed at room temperature; material was clear with no particulate matter. Added equal volume of MF59C adjuvant to thawed protein and mixed well by inverting the tube. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal. Used material within 1 hour of the addition of adjuvant. Needles were used for injections.
4	**Immunization 3-4: Protein immunization** The stock concentration was 1mg/ml. Immediately before immunizations, Iscomatrix was diluted to 0.1mg/ml with PBS (CFU U21). Added equal volume of 0.1mg/ml Iscomatrix into the 0.1mg/ml protein and mixed well. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal.
7-8, 10	**Immunization 1--3: Subtype B/C plasmid DNA in Saline** The immunogen was provided at 2.0mg/ml total DNA in sterile 0.9% saline. Stored at –80°C until use. Thawed DNA at room temperature; the material was clear or slightly opaque, with no particulate matter. Immunized each rabbit with 0.5ml DNA mixture per side (IM/Quadriceps), total 2 sides with 1.0ml per animal. Animals were shaved prior to immunization, under sedation of 1x dose IP (by animal weight) of Ketamine-Xylazine (80mg/ml - 4mg/ml). DNA injection used needle. Following the DNA injection, electroporation was administrated using a 6-needle circular array with 1cm diameter, 1cm needle length. Electroporation pulses were given at 20V/mm, 50ms pulse length, 1 pulse/s. **Immunization 3-4: Protein Immunization** Protein doses were 50ug protein per animal. The initial protein was diluted to 0.100 mg/ml in citrate buffer. Stored at –80°C until use. Thawed at room temperature; material was clear with no particulate matter. Added equal volume of MF59C adjuvant to thawed protein and mixed well by inverting the tube. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal. Used material within 1 hour of the addition of adjuvant. Needles were used for injections.

Group	Preparation
9	**Immunization 1-3: Subtype B plasmid DNA in Saline** The immunogen was provided at 1.0mg/ml total DNA in sterile 0.9% saline. Stored at –80°C until use. Thawed DNA at room temperature; the material was clear or slightly opaque, with no particulate matter. Immunized each rabbit with 0.5ml DNA mixture per side (IM/Quadriceps), total 2 sides with 1.0ml per animal. Animals were shaved prior to immunization, under sedation of 1x dose IP (by animal weight) of Ketamine-Xylazine (80mg/ml - 4mg/ml). DNA injection used needle. Following the DNA injection, electroporation was administrated using a 6-needle circular array with 1cm diameter, 1cm needle length. Electroporation pulses were given at 20V/mm, 50ms pulse length, 1 pulse/s. **Immunization 3-4: Protein Immunization** Protein doses were 50ug each protein per animal, total 100ug. The initial protein was diluted to 0.200 mg/ml in citrate buffer. Stored at –80°C until use. Thawed at room temperature; material was clear with no particulate matter. Added equal volume of MF59C adjuvant to thawed protein and mixed well by inverting the tube. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal. Used material within 1 hour of the addition of adjuvant. Needles were used for injections.
11	**Immunization 1-3: Subtype B plasmid DNA in Saline** The immunogen was provided at 1.0mg/ml total DNA in sterile 0.9% saline. Stored at –80°C until use. Thawed DNA at room temperature; the material was clear or slightly opaque, with no particulate matter. Immunized each rabbit with 0.5ml DNA mixture per side (IM/Quadriceps), total 2 sides with 1.0ml per animal. Animals were shaved prior to immunization, under sedation of 1x dose IP (by animal weight) of Ketamine-Xylazine (80mg/ml - 4mg/ml). DNA injection used needle. Following the DNA injection, electroporation was administrated using a 6-needle circular array with 1cm diameter, 1cm needle length. Electroporation pulses were given at 20V/mm, 50ms pulse length, 1 pulse/s. **Immunization 3-4: Protein Immunization** Protein doses were 50ug protein per animal. The initial protein was diluted to 0.100 mg/ml in citrate buffer. Stored at –80°C until use. Thawed at room temperature; material was clear with no particulate matter. Added equal volume of MF59C adjuvant to thawed protein and mixed well by inverting the tube. Immunized each rabbit with 0.5ml adjuvanted protein per side, IM/Glut for a total of 1ml per animal. Used material within 1 hour of the addition of adjuvant. Needles were used for injections.

Group	Preparation

The immunization (Table 7) schedules were as follows:

Table 7

Imm'n: Weeks: Group	*1* 0	*2* 4	*3* 12	*4* 24
1	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
2	Gp140 dV2 SF162 + Iscomatrix	Gp140 dV2 SF162 + Iscomatrix	Gp140 dV2 SF162 + Iscomatrix	Gp140 dV2 SF162 + Iscomatrix
3	pCMV 140 dV2 SF162 DNA	pCMV 140 dV2 SF162 DNA	pCMV 140 dV2 SF162 DNA Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
4	pCMV 140 dV2 SF162 DNA	pCMV 140 dV2 SF162 DNA	pCMV 140 dV2 SF162 DNA Gp140 dV2 SF162 + Iscomatrix	Gp140 dV2 SF162 + Iscomatrix
5	Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C
6	PCMV 140 dV2 TV1 DNA	pCMV140 dV2 TV1 DNA	pCMV140 dV2 TV1 DNA Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C
7	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
8	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C

Imm'n: Weeks: Group	*1* 0	*2* 4	*3* 12	*4* 24
9	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA Gp140 dV2 SF162 + MF59C Gp140 dV2 TV1 + MF59C (100ug Prot.)	Gp140 dV2 SF162 + MF59C Gp140 dV2 TV1 + MF59C (100ug Prot.)
10	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA Gp140 dV2 SF162 + MF59C Gp140 dV2 TV1 + MF59C (50ug Prot.)	Gp140 dV2 SF162 + MF59C Gp140 dV2 TV1 + MF59C (50ug Prot.)
11	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA	pCMV 140 dV2 SF162 DNA + PCMV 140 dV2 TV1 DNA (1.0mg) Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
	Note: all DNA was 1.0mg each except group 11 used 0.5mg DNA each.	**Note: all proteins were 50ug each except group 10 used 25ug ea.**		

Table 7 (cont.)

Imm'n: Weeks: Group	*5* 41	*6* 56
1	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
2	Gp140 dV2 SF162 + Iscomatrix	Gp140 dV2 SF162 + Iscomatrix
3	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
4	Gp140 dV2 SF162 + Iscomatrix	Gp140 dV2 SF162 + Iscomatrix
5	Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C
6	Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C
7	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C

Imm'n: Weeks: Group	*5* 41	*6* 56
8	Gp140 dV2 TV1 + MF59C	Gp140 dV2 TV1 + MF59C
9	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
	Gp140 dV2 TV1 + MF59C (100ug Prot.)	Gp140 dV2 TV1 + MF59C (100ug Prot.)
10	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
	Gp140 dV2 TV1 + MF59C (50ug Prot.)	Gp140 dV2 TV1 + MF59C (50ug Prot.)
11	Gp140 dV2 SF162 + MF59C	Gp140 dV2 SF162 + MF59C
	Note: all DNA was 1.0mg each except group 11 used 0.5mg DNA each.	**Note: all proteins were 50ug each except group 10 used 25ug each.**

The bleeding (Table 8) schedules for all groups (A-F) were as follows:

Table 8

Bleed:	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Week:	0	2	6	8	12	14	16	24
Sample:	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum

Bleed:	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
Week:	26	28	41	43	45	56	58	60
Sample:	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum	Clotted Bld. for Serum

To evaluate the combination effects of subtype C (TV1) and subtype B (SF162) gp140dV2 DNAs and proteins for DNA prime/boost on the generation of neutralizing antibody activity against HIV strain SF162 (type B) the following comparisons were carried out.

Neutralizing antibody responses against PBMC-grown SF162 and TV1 HIV-1 strains were monitored in the sera collected from the immunized rabbits using the following assay conducted essentially as follows. Virus neutralization was measured in 5.25.EGFP.Luc.M7 (M7-luc) cells obtained from Dr. Nathaniel Landau (Salk Institute, San Diego, CA). The format of this assay was essentially the same as the MT-2 assay that has been described elsewhere (Montefiori, et al., *J. Clin Microbiol.*

26:231-235, 1988) except that virus infection was quantified by luciferase reporter gene expression using a commercial luciferase kit (Promega). All serum samples were heat-inactivated for 1 hour at 56°C prior to assay. The virus stocks of the HIV-1 islolates were generated in PBMC. Neutralizing antibody titers are reported as reciprocal serum dilution at which 50% luciferase activity was measured in test wells as compared to virus control wells. Values shown in Figures 4 and 5 are the geometric mean titers plus standard errors of the neutralization titers for each group of animals.

The results of the assays for the presence of neutralizing antibodies are presented in Figure 4 and Figure 5. In the figures, the following Immunization Groups correspond to the Groups in Table 4: B DNA + B prot; C DNA + B prot (Group 6); B+C DNA + B prot (Group 7); B+C DNA + C prot (Group 8); B+C DNA & prot (Group 9); B+C DNA & prot (1/2) (Group 10); and, B+C DNA (1/2) + C prot (Group 11).

In Figure 4, the first vertical bar of each group of three bars is neutralizing activity against HIV-1 SF-162 in prebleed rabbit serum (Figure 4, Prebleed), the second vertical bar is serum from a bleed two weeks after the third immunization (Figure 4, 2 wk post 3rd), and the third vertical bar is serum from a bleed two weeks after the fourth immunization (Figure 4, 2 wk post 4th).

Figure 4 summarizes data showing the neutralization titers against HIV-1 SF162 between the 7 groups described above. These results demonstrated that all groups showed strong neutralizing activity against the HIV-1 SF162 isolate. Further, neutralizing activity significantly increased at post 4th immunization compared to post 3rd immunizations. Priming and boosting with B gene and B protein (B DNA + B prot) showed a high titer, as did the C gene and B protein (C DNA + B prot). For the mixed (B+C) DNA prime and single protein boost, B protein gave a high boost to the mixed gene prime (B+C DNA + B prot) and a boost to the C protein (B+C DNA + C prot). For the mixed DNA prime and protein boost, half dose (50ug) of protein (B+C DNA & prot (1/2)) induced high neutralizing activity as did the full dose of 100ug protein (B+C DNA & prot). The mixed DNA prime and single protein boost with subtype C protein, the half- dose (1mg) DNA (B+C DNA + C prot) also gave neutralizing activity, as did the full-dose of 2mg DNA (B+C DNA (1/2) + C prot).

In Figure 5, the prebleed values for neutralizing activity against HIV-1 TV1 in prebleed rabbit serum were less than one log for each group of bars (Figure 5, Prebleed), the grey vertical bars for each group are serum from bleeds two weeks after the fourth immunization (Figure 5, 2 wk post 4^{th}).

Figure 5 summarizes data showing the neutralization titers against HIV-1 TV1 (South African subtype C) between the 7 groups described above. These results demonstrated that all groups showed neutralizing activity against HIV1 subtype C TV1 isolate (as expected, because no subtype C DNA or protein was used, the B DNA + B protein showed the lowest neutralizing activity). For the mismatched a single DNA prime and a single protein boost (C DNA + B prot), priming with C gene and boosting with B protein showed a high titer, as did the B gene and B protein (B DNA + B prot). For the mixed (B+C) DNA prime and single protein boost, use of either B (B+C DNA + B prot) and C (B+C DNA + C prot) proteins had a similar boosting effect. For the mixed DNA prime and protein boost, full dose of 100ug protein (B+C DNA & prot) induced high neutralizing activity, as did the half dose of 50ug protein (B+C DNA & prot (1/2)). The half-dose (1mg) DNA (B+C DNA (1/2) + C prot) also gave neutralizing activity, as did the full-dose of 2mg DNA (B+C DNA + C prot).

Comparison of the data presented in Figure 4 and Figure 5 supported the combination methods of the present invention for generating an immune response in a subject. Such a comparison showed that the combination of DNA derived from different subtypes primed broad responses to multiple strains from different subtypes. This may indicate the targeting common conserved epitopes. Further, use of a single subtype protein was sufficient to boost broad neutralizing responses when immunity was primed with multiple strains from different subtypes of DNA. The DNA priming maintained the native envelope structure. This can induce T cell responses in addition to the B cell response. Finally, these results demonstrated that use of lower doses of proteins mixture can also provide strong immune responses.

These studies demonstrated the usefulness of the compositions and methods of the invention to generate immune responses, in particular to generate broad and potent neutralizing activity against diverse HIV strains.

Example 5

Immunogenicity Study of E1- E3 Deleted, Replication Defective Ad-HIV Recombinant Versus E3 Deleted, Replication Competent Ad-HIV Recombinant

The following experiments were carried out in chimpanzees. Chimpanzees with minimal Ad5- and Ad7-cross-reactive antibodies were selected for this experiment. Ad5 and Ad7 microtiter neutralization assays were performed essentially as previously described in Buge, et al., J. Virol. 71:8531-8541 (1997) and Lubeck, et al., Nature Med. 3:651-8 (1997). Chimpanzees were immunized according to the schedule in Table 9. Each group comprised 2 or 3 animals.

Table 9

Group	Number of Chimps	Week 0 (IN)	Week 12 (IN)	Week 36 (IM)
1	3	delAd5-E3-HIVgp160 10^7 pfu (replication competent)	Ad7delE3-HIVgp160 10^7 pfu (replication competent)	SF162 o-gp140V2 in MF59
2	2	Ad5delE3-HIVgp160 10^8 pfu (replication competent)	Ad7delE3-HIVgp160 10^8 pfu (replication competent)	SF162 o-gp140V2 in MF59
3	2	Ad5delE1/E3-HIVgp160 10^8 pfu (replication defective)	Ad7delE1/E3-HIVgp160 10^8 pfu (replication defective)	SF162 o-gp140V2 in MF59
4	3	Ad5delE1/E3-HIVgp160 10^9 pfu (replication defective)	Ad7delE1/E3-HIVgp160 10^9 pfu (replication defective)	SF162 o-gp140V2 in MF59

(IN = intranasal; IM = intramuscular)

The delAd5-E3, Ad7delE3, Ad5delE1/E3, and Ad7delE1/E3 vectors have been previously described (Nan X., et al., Development of an Ad7 cosmid system and generation of an Ad7deltaE1deltaE3HIV(MN) env/rev recombinant virus. Gene Ther. Feb;10(4):326-36 (2003)).

The Adeno-virus vectors (Ad recombinants) contained inserts derived from the HIV-1 subtype B prototype strain MN wherein the inserts encoded the gp160 envelope protein (see, e.g., GenBank Accession M17449; Gurgo,C., et al., "Envelope

sequences of two new United States HIV-1 isolates," Virology 164(2); 531-6 (1988); Lori, F., et al., "Effect of reciprocal complementation of two defective human immunodeficiency virus type 1 (HIV-1) molecular clones on HIV-1 cell tropism and virulence," J. Virol. 66(9); 5553-60 (1992); Lukashov, V.V., et al., "Increasing genotypic and phenotypic selection from the original genomic RNA populations of HIV-1 strains LAI and MN (NM) by peripheral blood mononuclear cell culture, B-cell-line propagation and T-cell-line adaptation," AIDS 9(12); 1307-11 (1995). HIV-1 MN is from one of the earliest available HIV-1 isolates, and is a commonly used reference and vaccine strain.

The MN isolate was taken from a six year old male pediatric AIDS patient from the area of Newark, New Jersey, USA in 1984. His mother was an IV drug user who died of pneumonia in 1982. His father was also HIV sero-positive. Other sequences from this patient from the 1984 blood sample and from a 1987 sample taken shortly before death (GenBank Accession U72495) are available. See also GenBank Accession L48364-L48379. The MN sequence was cloned from the isolate in lambda phage. The coding sequences for pol, nef and vpu are prematurely truncated; pol shows an in-frame stop codon at 3783, nef and vpu are prematurely truncated at position 9357 and position 6142 respectively. Another complete genome of the MN isolate is available with GenBank Accession number AF075719 and it too has defective genes; although not pol nor vpu. A set of V3 sequences from this isolate are available (GenBank Accession Accession numbers L48364-L48379; Lukashov, V. et al., AIDS 9:1307-1311 (1995)). The isolate MN is available from the NIH AIDS Reagent program, and is X4.

Ad-recombinant vectors (see Table 9) comprising HIV-1 MN gp160 protein coding sequences were diluted in PBS and administered drop-wise into the nostrils, 1 ml total volume, 500 µl per nostril. Antibiotics are administered for a total of 11 days, beginning 3 days prior to inoculation.

The polypeptide component used for a protein boost comprised SF162 o-gp140V2 protein. This protein is from the same HIV-1 subtype as the gp160 coding sequences used in the polynucleotide component, which were derived from HIV-1 MN. The SF162 o-gp140V2 protein was prepared using CMV3vector comprising the

gp140.mut7.mod.SF162.delV2 sequence expressed in CHO cells followed by oligo-protein isolation essentially as previously described, for example, in PCT International Publication No. WO/00/39302.

The protein boost was typically 100 ug of SF162 o-gp140V2 per chimpanzee. The SF162 o-gp140V2 protein was provided at 0.200 mg/ml in citrate buffer, stored at -80°C until use, and thawed at room temperature. The material as clear with no particulate matter. Equal volume of MF59C adjuvant was added. The mixture was stored at 4°C and mixed well by inverting the tube several times before use.

Each animal was immunized with a total volume of 1 ml per animal (using 1 or 2 IM sites per animal). Material was used within 1 hour of the addition of adjuvant.

Blood, secretory samples, and stool specimens were collected. Typically for blood samples, a 10 ml bleed was obtained for serum and a 30 ml bleed for heparinized blood.

The following assays were carried out on the collected samples.

A. Binding assays for HIV envelope antibodies by ELISA on immunized chimpanzee serum.

Standard HIV Env ELISA methods were employed in binding assays to detect HIV envelope antibodies in sera from chimpanzees immunized as just described. The methods were essentially as described by Buge, et al., J. Virol. 71:8531-8541 (1997) and Lubeck, et al., Nature Med. 3:651-8 (1997). Table 10 presents data for binding antibody titers to HIVIIIB gp160.

Table 10

Replication competent Ad	Dose	Chimp	At time of gp140 boost	2 weeks later	4 weeks later
	10^7	4x0363 (SW)	275	42,000	10,400
		4x0271 (SW)	7,500	21,800	10,700
		A163 (NI)	pending		
	10^8	4x0386	100	11,000	11,700

		(SW)			
		182D (NI)	100	800	410
Replication incompetent Ad	**Dose**	**Chimp**	**At time of gp140 boost**	**2 weeks later**	**4 weeks later**
	10^8	4x0376 (SW)	10	100	100
		4x0360 (SW)	pending		
	10^9	4x0373 (SW)	75	33,000	28,500
		87A003 (NI)	10	75	10
		A136 (NI)	10	590	440

Data for binding antibody titers to SF162 envelope protein was also evaluated and is shown in Figure 20.

B. Neutralizing antibody assays against TCLA and primary HIV isolates.

Virus neutralization against TCLA strains was measured in the MT-2 assay (Montefiori, et al., *J. Clin Microbiol.* 26:231-235 (1988)). Virus neutralization against primary HIV-1 strains was measured in M7-luc cells obtained from Dr. Nathaniel Landau (Salk Institute, San Diego, CA). The format of this assay was essentially the same as the MT-2 assay as described elsewhere (Montefiori, et al. *J. Clin Microbiol.* 26:231-235 (1988)) except that virus infection was quantified by luciferase reporter gene expression using a commercial luciferase kit (Promega). All serum samples were heat-inactivated for 1 hour at 56°C prior to assay. The virus stocks of the HIV-1 isolates were generated in PBMC.

Table 11 presents the neutralizing antibody data from these studies in chimpanzees.

Table 11

Group/Animal	**Vector/dose**	**Bleed day**	**HIV-1 MN[1]**	**HIV-1 SF162[2]**
1-1 4x0271 (SW)	delE3, 10^7	0	<20	<20
1-1		105	<20	<20

Group/Animal	Vector/dose	Bleed day	HIV-1 MN[1]	HIV-1 SF162[2]
1-1		273	48	40
1-2 4x0363 (SW)	delE3,10^7	0	<20	<20
1-2		105	25	92
1-2		273	1,296	5,877
2-1 4x0386 (SW)	delE3, 10^8	0	<20	<20
2-1		105	<20	20
2-1		273	228	133
2-2 182D (NI)	delE3, 10^8	0	<20	<20
2-2		105	47	97
2-2		273	5,801	3,437
3-1 4x0376 (SW)	delE1,E3, 10^8	0	<20	<20
3-1		105	<20	<20
3-1		273	<20	<20
4-1 4x0373 (SW)	delE1,E3, 10^9	0	<20	<20
4-1		105	34	<20
4-1		273	72	119
4-2 87A003 (NI)	delE1,E3, 10^9	0	<20	<20
4-2		105	<20	<20
4-2		273	<20	<20
4-3 A136 (NI)	delE1,E3, 10^9	0	<20	<20
4-3		105	<20	<20
4-3		273	<20	21

1 – determined by MT-2 assay described above. Neutralizing antibody titers are reported as reciprocal serum dilution at which 50% cell killing was measured in test wells as compared to virus control wells.

2 – determined by M7luc assay described above. Neutralizing antibody titers are reported as reciprocal serum dilution at which 50% luciferase activity was measured in test wells as compared to virus control wells.

The results in Table 11 support the use of the combination approaches described herein to induce potent and broad HIV-neutralization activity. For example,

on bleed day 273 sera obtain from all animals in Groups 1-3 comprised neutralizing antibodies against both the subtype B strain from which envelope protein coding sequences were derived (HIV-1 MN) for polynucleotide immunization and the subtype B strain from which envelope protein coding sequences were derived (HIV-1 SF162) for polypeptide immunization. One chimp in group 4 showed neutralizing antibodies, which correlated with better Ab response for that animal as seen in Table 10.

Overall, the replication competent recombinant Adeno vectors (delE3) provided a stronger priming of B cell responses than did the replication incompetent Adeno constructs (del E1, E3) with higher Env-specific binding antibody titers as measured by ELISA and higher serum neutralizing antibody responses against the MN and SF162 virus strains. These data demonstrate that a subject can be immunized with an envelope protein from a first HIV strain of a given subtype, be boosted with an envelope protein from a second HIV strain of the same subtype and generate neutralizing antibodies against both HIV strains. The data presented in Example 4 in combination with the data presented in Example 5 together demonstrate that such HIV strains may be within subtype, or from different subtypes.

Further assays may be used to evaluate the immune responses of the immunized chimpanzees, including, but not limited to, the following:

A. ELISPOT for Ad5 fiber, Ad7 fiber, HIV env, and HIVrev overlapping peptides.

Assay methods are essentially as described in Zhao, et al., J.Virol. 77:8354-8365 (2003). Peptides for use in this assay are derived from HIV-1MN Env and Rev or Ad fiber protein.

B. CTL assays by CR-release.

This standard CTL assay was carried out essentially as described by Lubeck, et al., Nature Med. 3:651-8 (1997), and Buge, et al., J. Virol. 71:8531-8541 (1997).

C. Proliferative T-cell responses against HIV gp120; aldrithiol-2 inactivated gp120; Ad proteins.

These assays were carried out essentially as described in Buge, et al., J. Virol. 71:8531-8541 (1997). The data is shown in Figure 21.

D. Ad5 and Ad7 microtiter neutralization assays.

These assays are carried out essentially as described in Buge, et al., J. Virol. 71:8531-8541 (1997).

E. Ad shedding in nasal and stool samples by PCR.

These assays are carried out essentially as described in Buge, et al., J. Virol. 71:8531-8541 (1997).

The data in this example support that the combination methods of the present invention can be used to broadly raise neutralizing antibodies against multiple viral strains of the same subtype.

Although preferred embodiments of the subject invention have been described in some detail, it is understood that obvious variations can be made without departing from the spirit and the scope of the invention. The following embodiments are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Exemplary Embodiments of the Present Invention:

1. A composition for generating an immune response in a mammal, said composition comprising,

a polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide derived from a first HIV strain of a first subtype, and

a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a second HIV strain of the first subtype, wherein said first HIV strain and said second HIV strain are different.

2. A composition for generating an immune response in a mammal, said composition comprising,

a polynucleotide component comprising two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides derived from a first HIV subtype, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains of the first subtype, and

a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that, if the polypeptide component comprises the same number or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain of the first subtype than the HIV immunogenic polypeptides provided by the polynucleotide component.

3. A composition for generating an immune response in a mammal, said composition comprising,

a polynucleotide component consisting essentially of one polynucleotide encoding an HIV immunogenic polypeptide derived from a first HIV strain of a first subtype, and

a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that at least one HIV immunogenic polypeptide of the polypeptide component is derived from a second HIV strain of the first subtype, wherein said first HIV strain and said second HIV strain are different;

with the provisos that (i) the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and (ii) the polypeptide component does not comprise an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

4. A composition for generating an immune response in a mammal, said composition comprising,

a polynucleotide component comprising two or more polynucleotide sequences comprising coding sequences for two or more analogous HIV immunogenic polypeptides derived from a first HIV subtype, wherein the coding sequences for at least two of the HIV immunogenic polypeptides are derived from different HIV strains of the first subtype, and

a polypeptide component comprising one or more HIV immunogenic polypeptides analogous to the polypeptide encoded by said polynucleotide component, with the proviso that, if the polypeptide component comprises the same number or greater than the number of analogous HIV immunogenic polypeptides encoded by the polynucleotide component, then at least one of the HIV immunogenic polypeptides of the polypeptide composition is derived from a different HIV strain of the first subtype than the HIV immunogenic polypeptides provided by the polynucleotide component;

with the further provisos that (i) the polynucleotide component does not encode an analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype, and (ii) the polypeptide component does not comprise an

analogous HIV immunogenic polypeptide derived from any subtype other than the first subtype.

5. The composition of any of embodiments 1 to 4, wherein said polynucleotide component comprises at least one polynucleotide that is a native polynucleotide.

6. The composition of any of embodiments 1 to 4, wherein said polynucleotide component comprises at least one polynucleotide that is a synthetic polynucleotide.

7. The composition of embodiment 6, wherein said synthetic polynucleotide comprises codons optimized for expression in mammalian cells.

8. The composition of embodiment 7, wherein said synthetic polynucleotide comprises codons optimized for expression in human cells.

9. The composition of any of embodiments 1 to 4, wherein the HIV immunogenic polypeptides are HIV envelope polypeptides.

10. The composition of embodiment 9, wherein at least one of said HIV polypeptides comprises one or more mutations.

11. The composition of embodiment 10, wherein at least one of said envelope polypeptides comprises a mutation in the cleavage site or a mutation in the glycosylation site.

12. The composition of embodiment 10, wherein at least one of said envelope polypeptides comprises a deletion or modification of the V1 region.

13. The composition of embodiment 10, wherein at least one of said envelope polypeptides comprises a deletion or modification of the V2 region.

14. The composition of embodiment 10, wherein at least one of said envelope polypeptides comprises a deletion or modification of the V3 region.

15. The composition of embodiment 10, wherein at least one of said envelope polypeptides comprises a deletion or modification of regions selected from the group consisting of the V1 region, the V2 region, the V3 region, and combinations thereof.

16. The composition of embodiment 10, wherein at least one of said envelope polypeptides comprises envelope polypeptide modified to expose an envelope binding region that binds to a CCR5 chemokine co-receptor.

17. The composition of any of embodiments 1 to 4, wherein at least one polynucleotide encoding an HIV immunogenic polypeptide encodes an immunogenic HIV polypeptide selected from the group consisting of: Gag, Env, Pol, Prot, Int, RT, vif, vpr, vpu, tat, rev, and nef.

18. The composition of any of embodiments 1 to 4, wherein the first HIV subtype is selected from the group consisting of: subtype A, subtype B, subtype C, subtype D, subtype E, subtype F, subtype G, subtype H, subtype I, subtype J, subtype K, subtype N and subtype O.

19. The composition of any of embodiments 1 to 4, wherein at least one of said immunogenic HIV polypeptides comprises one or more mutations.

20. The composition of any of embodiments 1 to 4, wherein said polynucleotide component further comprises a sequence encoding an additional antigenic polypeptide, with the proviso that the additional antigenic polypeptide is not an immunogenic polypeptide derived from an HIV-1 strain.

21. The composition of embodiment 20, wherein said polypeptide component further comprises a polypeptide having an additional antigenic peptide, with the proviso that the additional antigenic polypeptide is not an immunogenic polypeptide derived from an HIV-1 strain.

22. The composition of any of embodiments 1 to 4, wherein said polynucleotide component further comprises sequences encoding one or more control elements compatible with expression in a selected host cell, wherein said control elements are operable linked to polynucleotides encoding HIV immunogenic polypeptides.

23. The composition of embodiment 22, wherein said control elements are selected from the group consisting of a transcription promoter, a transcription enhancer element, a transcription termination signal, polyadenylation sequences, sequences for optimization of initiation of translation, an internal ribosome entry site, and translation termination sequences.

24. The composition of embodiment 23, wherein said transcription promoter is selected from the group consisting of CMV, CMV+intron A, SV40, RSV, HIV-Ltr, MMLV-ltr, and metallothionein.

25. A method of generating an immune response in a subject, comprising,

providing the composition for generating an immune response in a mammal of embodiments 1,2 3, or 4;

administering one or more gene delivery vectors comprising the polynucleotides of said polynucleotide component of the composition into said subject under conditions that are compatible with expression of said polynucleotides in said subject for the production of encoded HIV immunogenic polypeptides; and

administering the polypeptide component to said subject.

26. The method of embodiment 25, wherein said one or more gene delivery vectors and said polypeptide component are administered concurrently.

27. The method of embodiment 26, wherein said one or more gene delivery vectors and said polypeptide component are administered sequentially.

28. The method of embodiment 25, wherein said polypeptide component further comprises an adjuvant.

29. The method of embodiment 25, wherein said polynucleotide component further comprises a carrier.

30. The method of embodiment 25, wherein said one or more gene delivery vectors are nonviral vectors.

31. The method of embodiment 25, wherein said one or more gene delivery vectors are delivered using a particulate carrier.

32. The method of embodiment 31, wherein said one or more gene delivery vectors are coated on a gold or tungsten particle and said coated particle is delivered to said subject using a gene gun.

33. The method of embodiment 31, wherein said one or more gene delivery vectors are delivered using a PLG particle.

34. The method of embodiment 25, wherein said one or more gene delivery vectors are encapsulated in a liposome preparation.

35. The method of embodiment 25, wherein said one or more gene delivery vectors are viral vectors.

36. The method of embodiment 35, wherein said viral vectors are retroviral vectors.

37. The method of embodiment 35, wherein said viral vector are lentiviral vectors.

38. The method of embodiment 35, wherein said viral vectors are alphaviral vectors.

39. The method of embodiment 25, wherein said subject is a mammal.

40. The method of embodiment 39, wherein said mammal is a human.

41. The method of embodiment 25, wherein said immune response is a humoral immune response.

42. The method of embodiment 25, wherein said immune response is a cellular immune response.

43. The method of embodiment 25, wherein said one or more gene delivery vectors are administered intramuscularly, intramucosally, intranasally, subcutaneously, intradermally, transdermally, intravaginally, intrarectally, orally or intravenously.

44. The method of embodiment 25, wherein said immune response results in generating neutralizing antibodies in the subject against multiple strains derived from the first HIV subtype.

COMBINATION APPROACHES FOR GENERATING IMMUNE RESPONSES AGAINST MULTIPLE VIRAL STRAINS SELECTED FROM A GIVEN SUBTYPE OR SEROTYPE

Abstract of the Disclosure

The present invention relates to methods, polynucleotides, and polypeptides encoding immunogenic HIV polypeptides derived from different strains within an HIV subtype. Uses of the polynucleotides and polypeptides in combination approaches for generating immune responses are described. The combination approaches described herein have been shown to induce broad and potent neutralizing activity against diverse HIV strains from multiple strains within a given subtype. Formulations of compositions for generating immune responses and methods of use for such compositions are also disclosed.

FIGURE 1A

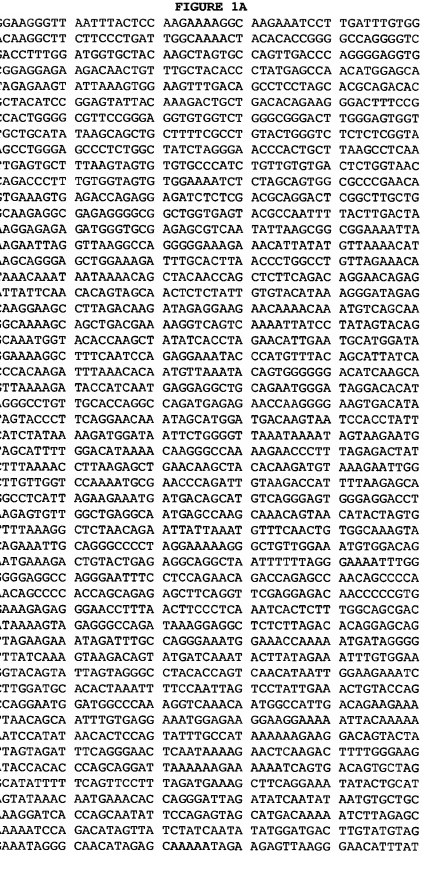

1 TGGAAGGGTT AATTTACTCC AAGAAAAGGC AAGAAATCCT TGATTTGTGG GTCTATCACA
61 CACAAGGCTT CTTCCCTGAT TGGCAAAACT ACACACCGGG GCCAGGGGTC AGATATCCAC
121 TGACCTTTGG ATGGTGCTAC AAGCTAGTGC CAGTTGACCC AGGGGAGGTG GAAGAGGCCA
181 ACGGAGGAGA AGACAACTGT TTGCTACACC CTATGAGCCA ACATGGACCA GAGGATGAAG
241 ATAGAGAAGT ATTAAAGTGG AAGTTTGACA GCCTCCTAGC ACGCAGACAC ATGGCCCGCG
301 AGCTACATCC GGAGTATTAC AAAGACTGCT GACACAGAAG GGACTTTCCG CCTGGGACTT
361 TCCACTGGGG CGTTCCGGGA GGTGTGGTCT GGGCGGGACT TGGGAGTGGT CAACCCTCAG
421 ATGCTGCATA TAAGCAGCTG CTTTTCGCCT GTACTGGGTC TCTCTCGGTA GACCAGATCT
481 GAGCCTGGGA GCCCTCTGGC TATCTAGGGA ACCCACTGCT TAAGCCTCAA TAAAGCTTGC
541 CTTGAGTGCT TTAAGTAGTG TGTGCCCATC TGTTGTGTGA CTCTGGTAAC TAGAGATCCC
601 TCAGACCCTT TGTGGTAGTG TGGAAAATCT CTAGCAGTGG CGCCCGAACA GGGACCAGAA
661 AGTGAAAGTG AGACCAGAGG AGATCTCTCG ACGCAGGACT CGGCTTGCTG AAGTGCACAC
721 GGCAAGAGGC GAGAGGGGCG GCTGGTGAGT ACGCCAATTT TACTTGACTA GCGGAGGCTA
781 GAAGGAGAGA GATGGGTGCG AGAGCGTCAA TATTAAGCGG CGGAAAATTA GATAAATGGG
841 AAAGAATTAG GTTAAGGCCA GGGGGAAAGA AACATTATAT GTTAAAACAT CTAGTATGGG
901 CAAGCAGGGA GCTGGAAAGA TTTGCACTTA ACCCTGGCCT GTTAGAAACA TCAGAAGGCT
961 GTAAACAAAT AATAAAACAG CTACAACCAG CTCTTCAGAC AGGAACAGAG GAACTTAGAT
1021 CATTATTCAA CACAGTAGCA ACTCTCTATT GTGTACATAA AGGGATAGAG GTACGAGACA
1081 CCAAGGAAGC CTTAGACAAG ATAGAGGAAG AACAAAACAA ATGTCAGCAA AAAGCACAAC
1141 AGGCAAAAGC AGCTGACGAA AAGGTCAGTC AAAATTATCC TATAGTACAG AATGCCCAAG
1201 GGCAAATGGT ACACCAAGCT ATATCACCTA GAACATTGAA TGCATGGATA AAAGTAATAG
1261 AGGAAAAGGC TTTCAATCCA GAGGAAATAC CCATGTTTAC AGCATTATCA GAAGGAGCCA
1321 CCCCACAAGA TTTAAACACA ATGTTAAATA CAGTGGGGGG ACATCAAGCA GCCATGCAAA
1381 TGTTAAAAGA TACCATCAAT GAGGAGGCTG CAGAATGGGA TAGGACACAT CCAGTACATG
1441 CAGGGCCTGT TGCACCAGGC CAGATGAGAG AACCAAGGGG AAGTGACATA GCAGGAACTA
1501 CTAGTACCCT TCAGGAACAA ATAGCATGGA TGACAAGTAA TCCACCTATT CCAGTAGAAG
1561 ACATCTATAA AAGATGGATA ATTCTGGGGT TAAATAAAAT AGTAAGAATG TATAGCCCTG
1621 TTAGCATTTT GGACATAAAA CAAGGGCCAA AAGAACCCTT TAGAGACTAT GTAGACCGGT
1681 TCTTTAAAAC CTTAAGAGCT GAACAAGCTA CACAAGATGT AAAGAATTGG ATGACAGACA
1741 CCTTGTTGGT CCAAAATGCG AACCCAGATT GTAAGACCAT TTTAAGAGCA TTAGGACCAG
1801 GGGCCTCATT AGAAGAAATG ATGACAGCAT GTCAGGGAGT GGGAGGACCT AGCCATAAAG
1861 CAAGAGTGTT GGCTGAGGCA ATGAGCCAAG CAAACAGTAA CATACTAGTG CAGAGAAGCA
1921 ATTTTAAAGG CTCTAACAGA ATTATTAAAT GTTTCAACTG TGGCAAAGTA GGGCACATAG
1981 CCAGAAATTG CAGGGCCCCT AGGAAAAAGG GCTGTTGGAA ATGTGGACAG GAAGGACACC
2041 AAATGAAAGA CTGTACTGAG AGGCAGGCTA ATTTTTTAGG GAAAATTTGG CCTTCCCACA
2101 AGGGGAGGCC AGGGAATTTC CTCCAGAACA GACCAGAGCC AACAGCCCCA CCAGCAGAAC
2161 CAACAGCCCC ACCAGCAGAG AGCTTCAGGT TCGAGGAGAC AACCCCCGTG CCGAGGAAGG
2221 AGAAAGAGAG GGAACCTTTA ACTTCCCTCA AATCACTCTT TGGCAGCGAC CCCTTGTCTC
2281 AATAAAAGTA GAGGGCCAGA TAAAGGAGGC TCTCTTAGAC ACAGGAGCAG ATGATACAGT
2341 ATTAGAAGAA ATAGATTTGC CAGGGAAATG GAAACCAAAA ATGATAGGGG GAATTGGAGG
2401 TTTTATCAAA GTAAGACAGT ATGATCAAAT ACTTATAGAA ATTTGTGGAA AAAAGGCTAT
2461 AGGTACAGTA TTAGTAGGGC CTACACCAGT CAACATAATT GGAAGAAATC TGTTAACTCA
2521 GCTTGGATGC ACACTAAATT TTCCAATTAG TCCTATTGAA ACTGTACCAG TAAAATTAAA
2581 ACCAGGAATG GATGGCCCAA AGGTCAAACA ATGGCCATTG ACAGAAGAAA AAATAAAAGC
2641 ATTAACAGCA ATTTGTGAGG AAATGGAGAA GGAAGGAAAA ATTACAAAAA TTGGGCCTGA
2701 TAATCCATAT AACACTCCAG TATTTGCCAT AAAAAAGAAG GACAGTACTA AGTGGAGAAA
2761 ATTAGTAGAT TTCAGGGAAC TCAATAAAAG AACTCAAGAC TTTTGGGAAG TTCAATTAGG
2821 AATACCACAC CCAGCAGGAT TAAAAAAGAA AAAATCAGTG ACAGTGCTAG ATGTGGGGGA
2881 TGCATATTTT TCAGTTCCTT TAGATGAAAG CTTCAGGAAA TATACTGCAT TCACCATACC
2941 TAGTATAAAC AATGAAACAC CAGGGATTAG ATATCAATAT AATGTGCTGC CACAGGGATG
3001 GAAAGGATCA CCAGCAATAT TCCAGAGTAG CATGACAAAA ATCTTAGAGC CCTTCAGAGC
3061 AAAAAATCCA GACATAGTTA TCTATCAATA TATGGATGAC TTGTATGTAG GATCTGACTT
3121 AGAAATAGGG CAACATAGAG CAAAAATAGA AGAGTTAAGG GAACATTTAT TGAAATGGGG

FIGURE 1B

```
3181 ATTTACAACA CCAGACAAGA AACATCAAAA AGAACCCCCA TTTCTTTGGA TGGGGTATGA
3241 ACTCCATCCT GACAAATGGA CAGTACAACC TATACTGCTG CCAGAAAAGG ATAGTTGGAC
3301 TGTCAATGAT ATACAGAAGT TAGTGGGAAA ATTAAACTGG GCAAGTCAGA TTTACCCAGG
3361 GATTAAAGTA AGGCAACTCT GTAAACTCCT CAGGGGGGCC AAAGCACTAA CAGACATAGT
3421 ACCACTAACT GAAGAAGCAG AATTAGAATT GGCAGAGAAC AGGGAAATTT TAAGAGAACC
3481 AGTACATGGA GTATATTATG ATCCATCAAA AGACTTGATA GCTGAAATAC AGAAACAGGG
3541 GCATGAACAA TGGACATATC AAATTTATCA AGAACCATTT AAAAATCTGA AAACAGGGAA
3601 GTATGCAAAA ATGAGGACTA CCCACACTAA TGATGTAAAA CAGTTAACAG AGGCAGTGCA
3661 AAAAATAGCC ATGGAAAGCA TAGTAATATG GGGAAAGACT CCTAAATTTA GACTACCCAT
3721 CCAAAAAGAA ACATGGGAGA CATGGTGGAC AGACTATTGG CAAGCCACCT GGATCCCTGA
3781 GTGGGAGTTT GTTAATACCC CTCCCCTAGT AAAATTATGG TACCAACTAG AAAAAGATCC
3841 CATAGCAGGA GTAGAAACTT TCTATGTAGA TGGAGCAACT AATAGGGAAG CTAAAATAGG
3901 AAAAGCAGGG TATGTTACTG ACAGAGGAAG GCAGAAAATT GTTACTCTAA CTAACACAAC
3961 AAATCAGAAG ACTGAGTTAC AAGCAATTCA GCTAGCTCTG CAGGATTCAG GATCAGAAGT
4021 AAACATAGTA ACAGACTCAC AGTATGCATT AGGAATCATT CAAGCACAAC CAGATAAGAG
4081 TGACTCAGAG ATATTTAACC AAATAATAGA ACAGTTAATA AACAAGGAAA GAATCTACCT
4141 GTCATGGGTA CCAGCACATA AAGGAATTGG GGGAAATGAA CAAGTAGATA AATTAGTAAG
4201 TAAGGGAATT AGGAAAGTGT TGTTTCTAGA TGGAATAGAT AAAGCTCAAG AAGAGCATGA
4261 AAGGTACCAC AGCAATTGGA GAGCAATGGC TAATGAGTTT AATCTGCCAC CCATAGTAGC
4321 AAAAGAAATA GTAGCTAGCT GTGATAAATG TCAGCTAAAA GGGGAAGCCA TACATGGACA
4381 AGTCGACTGT AGTCCAGGGA TATGGCAATT AGATTGTACC CATTTAGAGG GAAAAATCAT
4441 CCTGGTAGCA GTCCATGTAG CTAGTGGCTA CATGGAAGCA GAGGTTATCC CAGCAGAAAC
4501 AGGACAAGAA ACAGCATATT TTATATTAAA ATTAGCAGGA AGATGGCCAG TCAAAGTAAT
4561 ACATACAGAC AATGGCAGTA ATTTTACCAG TACTGCAGTT AAGGCAGCCT GTTGGTGGGC
4621 AGGTATCCAA CAGGAATTTG GAATTCCCTA CAATCCCCAA AGTCAGGGAG TGGTAGAATC
4681 CATGAATAAA GAATTAAAGA AAATAATAGG ACAAGTAAGA GATCAAGCTG AGCACCTTAA
4741 GACAGCAGTA CAAATGGCAG TATTCATTCA CAATTTTAAA AGAAAAGGGG GAATTGGGGG
4801 GTACAGTGCA GGGGAAAGAA TAATAGACAT AATAGCAACA GACATACAAA CTAAAGAATT
4861 ACAAAAACAA ATTATAAGAA TTCAAAATTT TCGGGTTTAT TACAGAGACA GCAGAGACCC
4921 TATTTGGAAA GGACCAGCCG AACTACTCTG GAAAGGTGAA GGGGTAGTAG TAATAGAAGA
4981 TAAAGGTGAC ATAAAGGTAG TACCAAGGAG GAAAGCAAAA ATCATTAGAG ATTATGGAAA
5041 ACAGATGGCA GGTGCTGATT GTGTGGCAGG TGGACAGGAT GAAGATTAGA GCATGGAATA
5101 GTTTAGTAAA GCACCATATG TATATATCAA GGAGAGCTAG TGGATGGGTC TACAGACATC
5161 ATTTTGAAAG CAGACATCCA AAAGTAAGTT CAGAAGTACA TATCCCATTA GGGGATGCTA
5221 GATTAGTAAT AAAAACATAT TGGGGTTTGC AGACAGGAGA AAGAGATTGG CATTTGGGTC
5281 ATGGAGTCTC CATAGAATGG AGACTGAGAG AATACAGCAC ACAAGTAGAC CCTGACCTGG
5341 CAGACCAGCT AATTCACATG CATTATTTTG ATTGTTTTAC AGAATCTGCC ATAAGACAAG
5401 CCATATTAGG ACACATAGTT TTTCCTAGGT GTGACTATCA AGCAGGACAT AAGAAGGTAG
5461 GATCTCTGCA ATACTTGGCA CTGACAGCAT TGATAAAACC AAAAAAGAGA AAGCCACCTC
5521 TGCCTAGTGT TAGAAAATTA GTAGAGGATA GATGGAACGA CCCCCAGAAG ACCAGGGGCC
5581 GCAGAGGGAA CCATACAATG AATGGACACT AGAGATTCTA GAAGAACTCA AGCAGGAAGC
5641 TGTCAGACAC TTTCCTAGAC CATGGCTCCA TAGCTTAGGA CAATATATCT ATGAAACCTA
5701 TGGGGATACT TGGACGGGAG TTGAAGCTAT AATAAGAGTA CTGCAACAAC TACTGTTCAT
5761 TCATTTCAGA ATTGGATGCC AACATAGCAG AATAGGCATC TTGCGACAGA GAAGAGCAAG
5821 AAATGGAGCC AGTAGATCCT AAACTAAAGC CCTGGAACCA TCCAGGAAGC CAACCTAAAA
5881 CAGCTTGTAA TAATTGCTTT TGCAAACACT GTAGCTATCA TTGTCTAGTT TGCTTTCAGA
5941 CAAAAGGTTT AGGCATTTCC TATGGCAGGA AGAAGCGGAG ACAGCGACGA AGCGCTCCTC
6001 CAAGTGGTGA AGATCATCAA AATCCTCTAT CAAAGCAGTA AGTACACATA GTAGATGTAA
6061 TGGTAAGTTT AAGTTTATTT AAAGGAGTAG ATTATAGATT AGGAGTAGGA GCATTGATAG
6121 TAGCACTAAT CATAGCAATA ATAGTGTGGA CCATAGCATA TATAGAATAT AGGAAATTGG
6181 TAAGACAAAA GAAAATAGAC TGGTTAATTA AAAGAATTAG GGAAAGAGCA GAAGACAGTG
6241 GCAATGAGAG TGATGGGGAC ACAGAAGAAT TGTCAACAAT GGTGGATATG GGGCATCTTA
6301 GGCTTCTGGA TGCTAATGAT TTGTAACACG GAGGACTTGT GGGTCACAGT CTACTATGGG
```

FIGURE 1C

6361 GTACCTGTGT GGAGAGAAGC AAAAACTACT CTATTCTGTG CATCAGATGC TAAAGCATAT
6421 GAGACAGAAG TGCATAATGT CTGGGCTACA CATGCTTGTG TACCCACAGA CCCCAACCCA
6481 CAAGAAATAG TTTTGGGAAA TGTAACAGAA AATTTTAATA TGTGGAAAAA TAACATGGCA
6541 GATCAGATGC ATGAGGATAT AATCAGTTTA TGGGATCAAA GCCTAAAGCC ATGTGTAAAG
6601 TTGACCCCAC TCTGTGTCAC TTTAAACTGT ACAGATACAA ATGTTACAGG TAATAGAACT
6661 GTTACAGGTA ATACAAATGA TACCAATATT GCAAATGCTA CATATAAGTA TGAAGAAATG
6721 AAAAATTGCT CTTTCAATGC AACCACAGAA TTAAGAGATA AGAAACATAA AGAGTATGCA
6781 CTCTTTTATA AACTTGATAT AGTACCACTT AATGAAAATA GTAACAACTT TACATATAGA
6841 TTAATAAATT GCAATACCTC AACCATAACA CAAGCCTGTC CAAAGGTCTC TTTTGACCCG
6901 ATTCCTATAC ATTACTGTGC TCCAGCTGAT TATGCGATTC TAAAGTGTAA TAATAAGACA
6961 TTCAATGGGA CAGGACCATG TTATAATGTC AGCACAGTAC AATGTACACA TGGAATTAAG
7021 CCAGTGGTAT CAACTCAACT ACTGTTAAAT GGTAGTCTAG CAGAAGAAGG GATAATAATT
7081 AGATCTGAAA ATTTGACAGA GAATACCAAA ACAATAATAG TACATCTTAA TGAATCTGTA
7141 GAGATTAATT GTACAAGGCC CAACAATAAT ACAAGGAAAA GTGTAAGGAT AGGACCAGGA
7201 CAAGCATTCT ATGCAACAAA TGACGTAATA GGAAACATAA GACAAGCACA TTGTAACATT
7261 AGTACAGATA GATGGAATAA AACTTTACAA CAGGTAATGA AAAAATTAGG AGAGCATTTC
7321 CCTAATAAAA CAATAAAATT TGAACCACAT GCAGGAGGGG ATCTAGAAAT TACAATGCAT
7381 AGCTTTAATT GTAGAGGAGA ATTTTTCTAT TGCAATACAT CAAACCTGTT TAATAGTACA
7441 TACTACCCTA AGAATGGTAC ATACAAATAC AATGGTAATT CAAGCTTACC CATCACACTC
7501 CAATGCAAAA TAAAACAAAT TGTACGCATG TGGCAAGGGG TAGGACAAGC AATGTATGCC
7561 CCTCCCATTG CAGGAAACAT AACATGTAGA TCAAACATCA CAGGAATACT ATTGACACGT
7621 GATGGGGGAT TTAACAACAC AAACAACGAC ACAGAGGAGA CATTCAGACC TGGAGGAGGA
7681 GATATGAGGG ATAACTGGAG AAGTGAATTA TATAAATATA AAGTGGTAGA AATTAAGCCA
7741 TTGGGAATAG CACCCACTAA GGCAAAAAGA AGAGTGGTGC AGAGAAAAAA AAGAGCAGTG
7801 GGAATAGGAG CTGTGTTCCT TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG
7861 TCAATAACGC TGACGGTACA GGCCAGACAA CTGTTGTCTG GTATAGTGCA ACAGCAAAGC
7921 AATTTGCTGA AGGCTATAGA GGCGCAACAG CATATGTTGC AACTCACAGT CTGGGGCATT
7981 AAGCAGCTCC AGGCGAGAGT CCTGGCTATA GAAAGATACC TAAAGGATCA ACAGCTCCTA
8041 GGGATTTGGG GCTGCTCTGG AAGACTCATC TGCACCACTG CTGTGCCTTG GAACTCCAGT
8101 TGGAGTAATA AATCTGAAGC AGATATTTGG GATAACATGA CTTGGATGCA GTGGGATAGA
8161 GAAATTAATA ATTACACAGA AACAATATTC AGGTTGCTTG AAGACTCGCA AAACCAGCAG
8221 GAAAAGAATG AAAAAGATTT ATTAGAATTG GACAAGTGGA ATAATCTGTG GAATTGGTTT
8281 GACATATCAA ACTGGCTGTG GTATATAAAA ATATTCATAA TGATAGTAGG AGGCTTGATA
8341 GGTTTAAGAA TAATTTTTGC TGTGCTCTCT ATAGTGAATA GAGTTAGGCA GGGATACTCA
8401 CCTTTGTCAT TTCAGACCCT TACCCCAAGC CCGAGGGGAC TGCAGTGGAA GGCAGTGATC
8461 GAAGAAGAAG GTGGAGAGCA AGACAGAGAC AGATCCATAC GATTGGTGAG CGGATTCTTG
8521 TCGCTTGCCT GGGACGATCT GCGGAGCCTG TGCCTCTTCA GCTACCACCG CTTGAGAGAC
8581 TTCATATTAA TTGCAGTGAG GGCAGTGGAA CTTCTGGGAC ACAGCAGTCT CAGGGGACTA
8641 CAGAGGGGGT GGGAGATCCT TAAGTATCTG GGAAGTCTTG TGCAGTATTG GGGTCTAGAG
8701 CTAAAAAAGA GTGCTATTAG TCCGCTTGAT ACCATAGCAA TAGCAGTAGC TGAAGGAACA
8761 GATAGGATTA TAGAATTGGT ACAAAGAATT TGTAGAGCTA TCCTCAACAT ACCTAGGAGA
8821 ATAAGACAGG GCTTTGAAGC AGCTTTGCTA TAAAATGGGA GGCAAGTGGT CAAAACGCAG
8881 CATAGTTGGA TGGCCTGCAG TAAGAGAAAG AATGAGAAGA ACTGAGCCAG CAGCAGAGGG
8941 AGTAGGAGCA GCGTCTCAAG ACTTAGATAG ACATGGGGCA CTTACAAGCA GCAACACACC
9001 TGCTACTAAT GAAGCTTGTG CCTGGCTGCA AGCACAAGAG GAGGACGGAG ATGTAGGCTT
9061 TCCAGTCAGA CCTCAGGTAC CTTTAAGACC AATGACTTAT AAGAGTGCAG TAGATCTCAG
9121 CTTCTTTTTA AAAGAAAAGG GGGGACTGGA AGGGTTAATT TACTCTAGGA AAAGGCAAGA
9181 AATCCTTGAT TTGTGGGTCT ATAACACACA AGGCTTCTTC CCTGATTGGC AAAACTACAC
9241 ATCGGGGCCA GGGGTCCGAT TCCCACTGAC CTTTGGATGG TGCTTCAAGC TAGTACCAGT
9301 TGACCCAAGG GAGGTGAAAG AGGCCAATGA AGGAGAAGAC AACTGTTTGC TACACCCTAT
9361 GAGCCAACAT GGAGCAGAGG ATGAAGATAG AGAAGTATTA AAGTGGAAGT TTGACAGCCT
9421 TCTAGCACAC AGACACATGG CCCGCGAGCT ACATCCGGAG TATTACAAAG ACTGCTGACA

FIGURE 1D

```
9481 CAGAAGGGAC TTTCCGCCTG GGACTTTCCA CTGGGGCGTT CCGGGAGGTG TGGTCTGGGC
9541 GGGACTTGGG AGTGGTCACC CTCAGATGCT GCATATAAGC AGCTGCTTTT CGCTTGTACT
9601 GGGTCTCTCT CGGTAGACCA GATCTGAGCC TGGGAGCTCT CTGGCTATCT AGGGAACCCA
9661 CTGCTTAGGC CTCAATAAAG CTTGCCTTGA GTGCTCTAAG TAGTGTGTGC CCATCTGTTG
9721 TGTGACTCTG GTAACTAGAG ATCCCTCAGA CCCTTTGTGG TAGTGTGGAA AATCTCTAGC
9781 A
```

Figure 2A

↓: indicates the regions for β-sheet and V1/V2 loop deletions

*: is the N-linked glycosylation sites for subtype C TV1 and TV2. Possible mutation (N→ Q) or deletions can be performed.

```
                    1                                                  50
      B-SF162   (1) ----MDAMKRGLCCVLLLCGAVFVSP-SAVEKLWVTVYYGVPVWKEATTT
     C-TV1.8_2  (1) MRVMGTQKNCQQWWIWGILGFWMLMI-CNTEDLWVTVYYGVPVWRDAKTT
     C-TV1.8_5  (1) MRVMGTQKNCQQWWIWGILGFWMLMI-CNTEDLWVTVYYGVPVWREAKTT
  C-TV2.12-5/1  (1) MRARGILKNYRHWWIWGILGFWMLMM-CNVKGLWVTVYYGVPVGREAKTT
        C-MJ4   (1) MRVKGIPRNWQQWWIWGSLGFWIIC--SVMGNLWVTVYYGVPVWREAKTT
IndiaC-93IN101  (1) MRVRGTLRNYQQWWIWGVLGFWMLMICNGGGNLWVTVYYGVPVWKEAKTT
      A-Q2317   (1) MRVMGIQRNCQHLLTWGIMILGTIIFCSAVENLWVTVYYGVPVWRDADTT
     D-92UG001  (1) MRVREIERNYLCLWRWGIMLLGMLMTYSVAEKKWVTVYYGVPVWKEATTT
      E-cm235   (1) ----MDAMKRGLCCVLLLCGAVFVSP-SASNNLWVTVYYGVPVWRDADTT
     Consensus  (1) MRV G  RN Q WWIWGILGFWMLM  S  E LWVTVYYGVPVWREAKTT

                   51                                  *              100
      B-SF162  (46) LFCASDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMV
     C-TV1.8_2 (50) LFCASDAKAYETEVHNVWATHACVPTDPNPQEIVLGNVTENFNMWKNDMA
     C-TV1.8_5 (50) LFCASDAKAYETEVHNVWATHACVPTDPNPQEIVLGNVTENFNMWKNNMA
  C-TV2.12-5/1 (50) LFCASDAKAYEKEVHNVWATHACVPTDPNPQEVILGNVTENFNMWKNDMV
        C-MJ4  (49) LFCASDAKAYEAEVHNVWATHACVPTDPNPQEIELKNVTENFNMWENDMV
IndiaC-93IN101 (51) LLCASDAKAYEREVHNVWATHACVPTDPNPQEIVLGHVTENFNMWKNDMV
      A-Q2317  (51) LFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLDNVTEKFNMWKNNMV
     D-92UG001 (51) LFCASDAKSYKTEVHNIWATHACVPTDPNPREIELENVTENFNMWKNNMV
      E-cm235  (46) LFCASDAKAHETEVHNVWATHACVPTDPNPQEIHLENVTENFNMWKNNMV
     Consensus (51) LFCASDAKAYETEVHNVWATHACVPTDPNPQEIVL NVTENFNMWKNNMV

                                                  β2/V1V2/β3
                   101                          ↓  *   *   *   *   *150
      B-SF162  (96) EQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATN--------TKSS
     C-TV1.8_2 (100) DQMHEDVISLWDQSLKPCVKLTPLCVTLNCTDTNVTGNRTVTGNSTNNTN
     C-TV1.8_5 (100) DQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDTNVTGNRTVTGNTNDTNI
  C-TV2.12-5/1 (100) DQMQEDIISLWDQSLKPCVKLTPLCVTLNCTNATVNYN-------NTS--
        C-MJ4  (99) DQMHEDIISLWDQSLKPCVKLTPLCVTLNCKNVTSKDI----------NI
IndiaC-93IN101 (101) DQMHEDVISLWDQSLKPCVKLTPLCVTLECRNVSRNVS----------SY
      A-Q2317  (101) EQMHTDIISLWDQSLKPCVKLTPLCVTLHCTNVTSVNT----------
     D-92UG001 (101) EQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDARRNET-------RNNIT
      E-cm235  (96) EQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNAKLTNV-----NNITSVS
     Consensus (101) DQMHEDIISLWDQSLKPCVKLTPLCVTLNCTN

                   151       *   *                                  200
      B-SF162 (138) NWKEMDRGEIKNCSFKVTTSIRNKMQKEYALFYKLDVVPIDN-----DNT
     C-TV1.8_2 (150) GTGIYNIEEMKNCSFNATTELRDKKHKEYALFYRLDIVPLN---ENSDNF
     C-TV1.8_5 (150) ANATYKYEEMKNCSFNATTELRDKKHKEYALFYKLDIVPLN---ENSNNF
  C-TV2.12-5/1 (141) --------KDMKNCSFYVTTELRDKKKENALFYRLDIVPLNNR-KNGNIN
        C-MJ4  (139) TSNAEMKAEMKNCSFNITTELRDKKQKEYALFYKLDIVPLTNDNASENAS
IndiaC-93IN101 (141) NTYNGSVEEIKNCSFNATPEVRDRKQRMYALFYGLDIVPLNKKNSSENSS
      A-Q2317  (139) -TG--DREGLKNCSFNMTTELRDKRQKVYSLFYRLDIVPINE----NQGS
     D-92UG001 (144) GMENNDQIEMKNCSFNITTKLIDKKKQVHALFYRLDVVQIDNDTSNSNYS
      E-cm235  (141) NTIGNITDEVRNCSFNMTTELRDKKQKVHALFYKLDIVPIED---NKTSS
     Consensus (151)  T     EEMKNCSFNITTELRDKK KEYALFYKLDIVPLN    N N S
```

Figure 2B

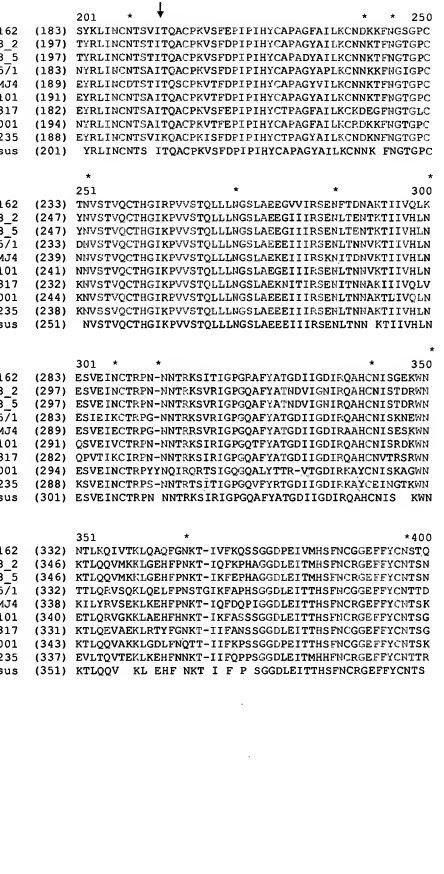
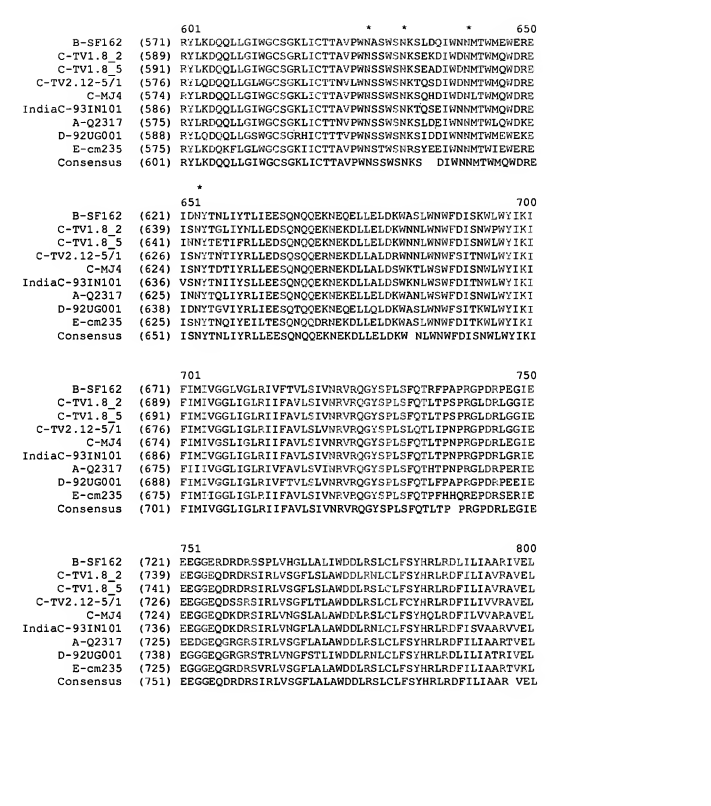

201 * 250
B-SF162 (183) SYKLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGSGPC
C-TV1.8_2 (197) TYRLINCNTSTITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPC
C-TV1.8_5 (197) TYRLINCNTSTITQACPKVSFDPIPIHYCAPADYAILKCNNKTFNGTGPC
C-TV2.12-571 (183) NYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAPLKCNNKKFNGIGPC
C-MJ4 (189) EYRLINCDTSTITQSCPKVTFDPIPIHYCAPAGYVILKCNNKTFNGTGPC
IndiaC-93IN101 (191) EYRLINCNTSAITQACPKVTFDPIPIHYCAPAGYAILKCNNKTFNGTGPC
A-Q2317 (182) EYRLINCNTSAITQACPKVSFEPIPIHYCTPAGFAILKCKDEGFNGTGLC
D-92UG001 (194) NYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCPDKKFNGTGPC
E-cm235 (188) EYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPC
Consensus (201) YRLINCNTS ITQACPKVSFDPIPIHYCAPAGYAILKCNNK FNGTGPC
251 300
B-SF162 (233) TNVSTVQCTHGIRPVVSTQLLLNGSLAEEGVVIRSENFTDNAKTIIVQLK
C-TV1.8_2 (247) YNVSTVQCTHGIKPVVSTQLLLNGSLAEEGIIIRSENLTENTKTIIVHLN
C-TV1.8_5 (247) YNVSTVQCTHGIKPVVSTQLLLNGSLAEEGIIIRSENLTENTKTIIVHLN
C-TV2.12-571 (233) DNVSTVQCTHGIKPVVSTQLLLNGSLAEEIIIRSENLTNNVKTIIVHLN
C-MJ4 (239) NNVSTVQCTHGIKPVVSTQLLLNGSLAEKEIIIRSKNITDNVKTIIVHLN
IndiaC-93IN101 (241) NNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENLTNHVKTIIVHLN
A-Q2317 (232) KNVSTVQCTHGIKPVVSTQLLLNGSLAEKNITIRSENITNNAKIIIVQLV
D-92UG001 (244) KNVSTVQCTHGIRPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTLIVQLN
E-cm235 (238) KNVSSVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLN
Consensus (251) NVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNN KTIIVHLN
301 * * * 350
B-SF162 (283) ESVEINCTRPN-NNTRKSITIGPGRAFYATGDIIGDIRQAHCNISGEKWN
C-TV1.8_2 (297) ESVEINCTRPN-NNTRKSVRIGPGQAFYATNDVIGNIRQAHCNISTDRWN
C-TV1.8_5 (297) ESVEINCTRPN-NNTRKSVRIGPGQAFYATNDVIGNIRQAHCNISTDRWN
C-TV2.12-571 (283) ESIEIKCTRPG-NNTRKSVRIGPGQAFYATGDIIGDIRQAHCNISKNEWN
C-MJ4 (289) ESVEIECTRPG-NNTRRSVRIGPGQAFYATGDIIGDIRAAHCNISESKWN
IndiaC-93IN101 (291) QSVEIVCTRPN-NNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISRDKWN
A-Q2317 (282) QPVTIKCIRPN-NNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVTRSRWN
D-92UG001 (294) ESVEINCTRPYYNQIRQRTSIGQGQALYTTR-VTGDIRKAYCNISKAGWN
E-cm235 (288) KSVEINCTRPS-NNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWN
Consensus (301) ESVEINCTRPN NNTRKSIRIGPGQAFYATGDIIGDIRQAHCNIS KWN
351 *400
B-SF162 (332) NTLKQIVTKLQAQFGNKT-IVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQ
C-TV1.8_2 (346) KTLQQVMKKLGEHFPNKT-IQFKPHAGGDLEITMHSFNCRGEFFYCNTSN
C-TV1.8_5 (346) KTLQQVMKKLGEHFPNKT-IKFEPHAGGDLEITMHSFNCRGEFFYCNTSN
C-TV2.12-571 (332) TTLQRVSQKLQELFPNSTGIKFAPHSGGDLEITTHSFNCGGEFFYCNTTD
C-MJ4 (338) KILYRVSEKLKEHFPNKT-IQFDQPIGGDLEITTHSFNCRGEFFYCNTSK
IndiaC-93IN101 (340) ETLQRVGKKLAEHFHNKT-IKFASSSGGDLEITTHSFNCRGEFFYCNTSG
A-Q2317 (331) KTLQEVAEKLRTYFGNKT-IIFANSSGGDLEITTHSFNCGGEFFYCNTSG
D-92UG001 (343) KTLQQVAKKLGDLFNQTT-IIFKPSSGGDPEITTHSFNCGGEFFYCNTSK
E-cm235 (337) EVLTQVTEKLKEHFNNKT-IIFQPPSGGDLEITMHHFNCRGEFFYCNTTR
Consensus (351) KTLQQV KL EHF NKT I F P SGGDLEITTHSFNCRGEFFYCNTS

Figure 2C


401
β20/β21
450
451
500
501
550
551
600
B-SF162 (381) LFNSTWNN-----TIGPNN--TNGTITLPCRIKQIINRWQEVGKAMYAPP
C-TV1.8_2 (395) LFNSTYHS---NNGTYKYNGNSSSPITLQCKIKQIVRMWQGVGQATYAPP
C-TV1.8_5 (395) LFNSTYYP---KNGTYKYNGNSSLPITLQCKIKQIVRMWQGVGQAMYAPP
C-TV2.12-571 (382) LFNSTYSNGTCTNGTCMSN--NTERITLQCRIKQIINMWQEVGRAMYAPP
C-MJ4 (387) LFNGTYNS-----TGDTSN----STITLSCRIKQIINMWQGVGRAMYASP
IndiaC-93IN101 (389) LFNGTYMPTYMPNGTESNS---NSTITIPCRIKQIINMWQEVGRAMYAPP
A-Q2317 (380) LFNSTWYVNSTWNDTDSTQ-ESNDTITLPCRIKQIINMWQRAGQAMYAPP
D-92UG001 (392) LFNSAWND-STWNIGNNNTGSDNETIIIPCRIKQIINMWQGVGKAMYAPP
E-cm235 (386) LFNNTCIE----NGTMGGC---NGTIILPCKIKQIINMWQGAGQAMYAPP
Consensus (401) LFNSTY NGT N N TITLPCRIKQIINMWQGVGRAMYAPP
B-SF162 (424) IRGQIRCSSNITGLLLTRDGGKEISNT---TEIFRPGGGDMRDNWRSELY
C-TV1.8_2 (442) IAGNITCRSNITGILLTRDGGFNTTNN---TETFRPGGGDMRDNWRSELY
C-TV1.8_5 (442) IAGNITCRSNITGILLTRDGGFNNTNNDT-EETFRPGGGDMRDNWRSELY
C-TV2.12-571 (430) IAGNITCRSNITGLLLTRDGGDNNTET----ETFRPGGGDMRDNWRSELY
C-MJ4 (428) IAGNITCKSNITGLLLTRDGGNETSGI----EIFRPAGGDMRDNWRSELY
IndiaC-93IN101 (436) IAGNITCTSNITGLLLVHDGGIKENDTENKTEIFRPGGGDMRDNWRSELY
A-Q2317 (429) IPGVIKCESNITGLLLTRDGGKDNNVN----ETFRPGGGDMRDNWRSELY
D-92UG001 (441) IEGWINCASNITGLLLVRDGGGANDSQ---NETFRPQGGDMRDNWRSELY
E-cm235 (429) ISGRINCVSNITGILLTRDGGAINTTN----ETFRPGGGNIKDNWPSELY
Consensus (451) IAGNITC SNITGLLLTRDGG NT N ETFRPGGGDMRDNWRSELY
B-SF162 (471) KYKVVKIEPLGVAPTKAKRRVVQREKRAVTLGAMFLGFLGAAGSTMGARS
C-TV1.8_2 (489) KYKVVEIKPLGIAPTKAKRRVVQREKRAVGIGAVFLGFLGAAGSTMGAAS
C-TV1.8_5 (491) KYKVVEIKPLGIAPTKAKRRVVQRKKRAVGIGAVFLGFLGAAGSTMGAAS
C-TV2.12-571 (476) KYKVVEIKPLGVAPTAAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAAS
C-MJ4 (474) KYKVVEIKPLGLAPTKSKRRVVEREKRAVTFGAMFLGFLGAAGSTMGAAS
IndiaC-93IN101 (486) KYKVVEIKPLGVAPTAAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAAS
A-Q2317 (475) KYKVVEIEPLGVAPTRAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGATS
D-92UG001 (488) KYKVVKIEPLGIAPTKAKRRVVEREKRAIGLGAMFLGFLGAAGSTMGAAS
E-cm235 (475) KYKVVQIEPLGIAPTRAKRRVVEREKRAVGIGAMIFGFLGAAGSTMGAAS
Consensus (501) KYKVVEIKPLGIAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAAS
B-SF162 (521) LTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVE
C-TV1.8_2 (539) ITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLQLTVWGIKQLQARVLAIE
C-TV1.8_5 (541) ITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLQLTVWGIKQLQARVLAIE
C-TV2.12-571 (526) ITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIE
C-MJ4 (524) MTLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRVLAVE
IndiaC-93IN101 (536) ITLTAQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQTRVLAIE
A-Q2317 (525) ITLTVQARQLLSGIVQQQNNLLRAIEAQQHLLKLTVWGIKQLQARVLAVE
D-92UG001 (538) LTLTVQARQLLSGIVQHQNNLLMAIEAQQHLLQLTVWGIKQLQARILAVE
E-cm235 (525) ITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVE
Consensus (551) ITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVE

Figure 2D


601 * * * 650
B-SF162 (571) RYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLDQIWNNMTWMEWERE
C-TV1.8_2 (589) RYLKDQQLLGIWGCSGRLICTTAVPWNSSWSNKSEKDIWDNMTWMQWDRE
C-TV1.8_5 (591) RYLKDQQLLGIWGCSGRLICTTAVPWNSSWSNKSEADIWDNMTWMQWDRE
C-TV2.12-571 (576) RYLQDQQLLGLWGCSGKLICTTNVLWNSSWSNKTQSDIWDNMTWMQWDRE
C-MJ4 (574) RYLRDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQHDIWDNLTWMQWDRE
IndiaC-93IN101 (586) RYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKTQSEIWNNMTWMQWDRE
A-Q2317 (575) RYLRDQQLLGIWGCSGKLICTTNVPWNSSWSNKSLDEIWNNMTWLQWDKE
D-92UG001 (588) RYLQDQQLLGSWGCSGRHICTTTVPWNSSWSNKSIDDIWNNMTWMEWEKE
E-cm235 (575) RYLKDQKFLGLWGCSGKIICTTAVPWNSTWSNRSYEEIWNNMTWIEWERE
Consensus (601) RYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKS DIWNNMTWMQWDRE

651 700
B-SF162 (621) IDNYTNLIYTLIEESQNQQEKNEQELLELDKWASLWNWFDISKWLWYIKI
C-TV1.8_2 (639) ISNYTGLIYNLLEDSQNQQEKNEKDLLELDKWNNLWNWFDISNWPWYIKI
C-TV1.8_5 (641) INNYTETIFRLLEDSQNQQEKNEKDLLELDKWNNLWNWFDISNWLWYIKI
C-TV2.12-571 (626) ISNYTNTIYRLLEDSQSQQERNEKDLLALDRWNNLWNWFSITNWLWYIKI
C-MJ4 (624) ISNYTDTIYRLLEESQNQQERNEKDLLALDSWKTLWSWFDISNWLWYIKI
IndiaC-93IN101 (636) VSNYTNIIYSLLEESQNQQEKNEKDLLALDSWKNLWSWFDITNWLWYIKI
A-Q2317 (625) INNYTQLIYRLIEESQNQQEKNEKELLELDKWANLWSWFDISNWLWYIKI
D-92UG001 (638) IDNYTGVIYRLIEESQTQQEKNEQELLQLDKWASLWNWFSITKWLWYIKI
E-cm235 (625) ISNYTNQIYEILTESQNQQDRNEKDLLELDKWASLWNWFDITKWLWYIKI
Consensus (651) ISNYTNLIYRLLEESQNQQEKNEKDLLELDKW NLWNWFDISNWLWYIKI

701 750
B-SF162 (671) FIMIVGGLVGLRIVFTVLSIVNRVRQGYSPLSFQTRFPAPRGPDRPEGIE
C-TV1.8_2 (689) FIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPSPRGLDRLGGIE
C-TV1.8_5 (691) FIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPSPRGLDRLGGIE
C-TV2.12-571 (676) FIMIVGGLIGLRIIFAVLSLVNRVRQGYSPLSLQTLIPNPRGPDRLGGIE
C-MJ4 (674) FIMIVGSLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPNPRGLDRLEGIE
IndiaC-93IN101 (686) FIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPNPRGPDRLGRIE
A-Q2317 (675) FIIIVGGLIGLRIVFAVLSVINRVRQGYSPLSFQTHTPNPRGLDRPERIE
D-92UG001 (688) FIMIVGGLIGLRIVFTVLSLVNRVRQGYSPLSFQTLFPAPRGPDRPEEIE
E-cm235 (675) FIMIIGGLIGLRIIFAVLSIVNEVRQGYSPLSFQTPFHHQREPDRSERIE
Consensus (701) FIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTP PRGPDRLEGIE

751 800
B-SF162 (721) EEGGERDRDRSSPLVHGLLALIWDDLRSLCLFSYHRLRDLILIAARIVEL
C-TV1.8_2 (739) EEGGEQDRDRSIRLVSGFLSLAWDDLRNLCLFSYHRLRDFILIAVRAVEL
C-TV1.8_5 (741) EEGGEQDRDRSIRLVSGFLSLAWDDLRSLCLFSYHRLRDFILIAVRAVEL
C-TV2.12-571 (726) EEGGEQDSSRSIRLVSGFLTLAWDDLRSLCLFCYHRLRDFILIVVRAVEL
C-MJ4 (724) EEGGEQDKDRSIRLVNGSLALAWDDLRSLCLFSYHQLRDFILVVARAVEL
IndiaC-93IN101 (736) EEGGEQDKDRSIRLVNGFLALAWDDLRNLCLFSYHRLRDFISVAARVVEL
A-Q2317 (725) EEDGEQGRGRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVEL
D-92UG001 (738) EGGGEQGRGRSTRLVNGFSTLIWDDLRNLCLFSYHRLRDLILIATRIVEL
E-cm235 (725) EGGGEQGRDRSVRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVKL
Consensus (751) EEGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAAR VEL

Figure 2E

```
                    801                                                 850
      B-SF162 (771) LGR-------RGWEALKYWGNLLQYWIQELKNSAVSLFDAIAIAVAEGTD
    C-TV1.8_2 (789) LGHSSLRGLQRGWEILKYLGSLVQYWGLELKKSAISLLDTIAITVAEGTD
    C-TV1.8_5 (791) LGHSSLRGLQRGWEILKYLGSLVQYWGLELKKSAISPLDTIAIAVAEGTD
  C-TV2.12-571 (776) LGHSSLRGLQRGWGTLKYLGSLVQYWGLELKKSAINLLDTIAIAVAEGTD
        C-MJ4 (774) LGRSSLRGLQRGWEALKYLGSLIQYWGLELKRSTISLLDTVAIAVAEGTD
IndiaC-93IN101 (786) LGRSS-------WEALKYLGSLVQYWGLELKKSAISLFDSIAIVVAEGTD
      A-Q2317 (775) LGHSSLKGLRLGWEGIKYLWNLLSYWGRELKISAINLVDTIAIAVAGWTD
    D-92UG001 (788) LGR-------RGWEAIKYLWNLLQYWSQELKTSAISLFNATAVAVAEGTD
      E-cm235 (775) LGRSSLKGLRRGWEGLKYLGHLLLYWGQELKISAISLLDATAIIVAGWTD
    Consensus (801) LGRSSLRGL RGWEALKYLGSLLQYWGLELKKSAISLLDTIAIAVAEGTD


                    851                         879
      B-SF162 (814) RIIEVAQRIGRAFLHIPRRIRQGFERALL
    C-TV1.8_2 (839) RIIELVQRICRAILNIPRRIRQGFEAALL
    C-TV1.8_5 (841) RIIELVQRICRAILNIPRRIPQGFEAALL
  C-TV2.12-571 (826) RILEFIQNLCRGIRNVPRRIRQGFEAALQ
        C-MJ4 (824) RIIELIQRIWRAICNIPRRIRQGFEAALL
IndiaC-93IN101 (829) RIIELVQGFCRAIRNIPTRIRQGFEAALQ
      A-Q2317 (825) RVIEIAQRIGRAILHIPVRIRQGLERALL
    D-92UG001 (831) RVIEVVQRFRGILNVPTRIPQGLERALL
      E-cm235 (825) RVIEVAQGAWRAILHIPRRIRQGLERTLL
    Consensus (851) RIIELVQRI RAILNIPRRIRQGFEAALL
```

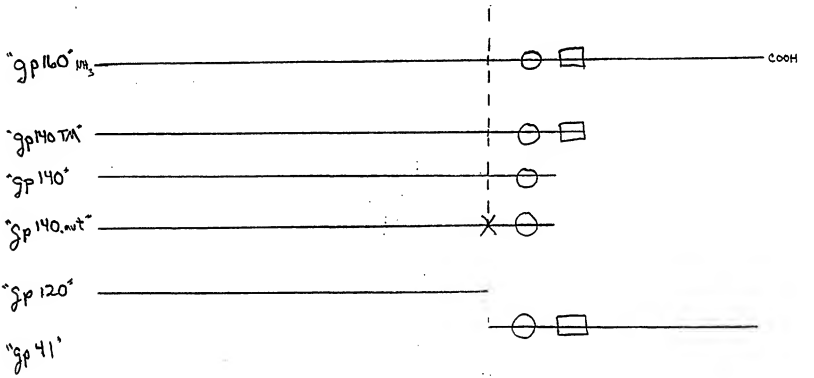


Figure 3

Figure 4

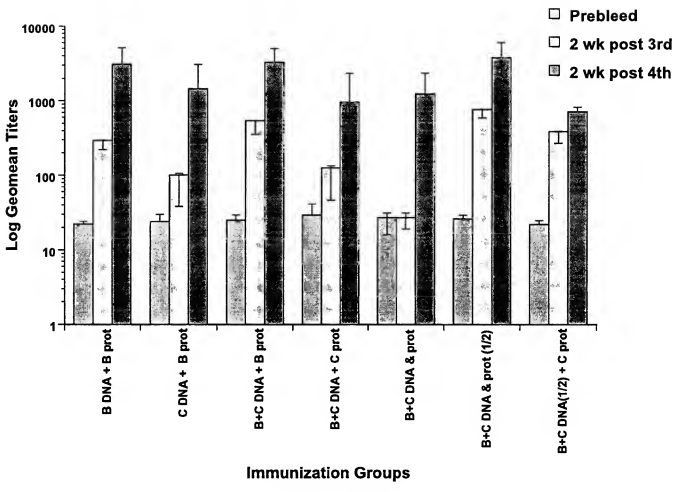
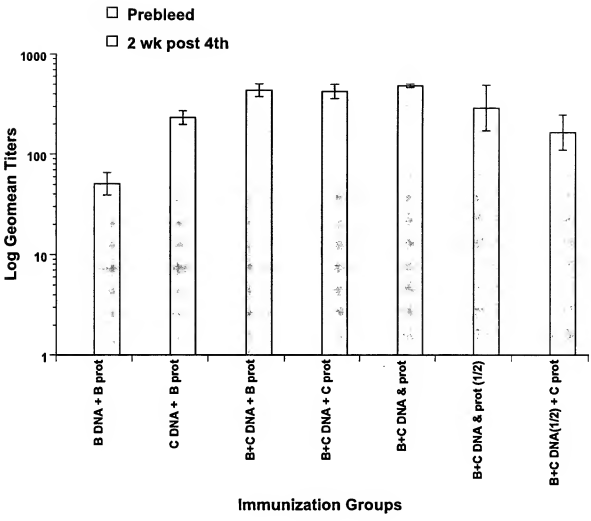

Prebleed
2 wk post 3rd
2 wk post 4th
Log Geomean Titers
10000
1000
100
10
1
B DNA + B prot
C DNA + B prot
B+C DNA + B prot
B+C DNA + C prot
B+C DNA & prot
B+C DNA & prot (1/2)
B+C DNA(1/2) + C prot
Immunization Groups

Figure 5


Prebleed
2 wk post 4th
Log Geomean Titers
1000
100
10
1
B DNA + B prot
C DNA + B prot
B+C DNA + B prot
B+C DNA + C prot
B+C DNA & prot
B+C DNA & prot (1/2)
B+C DNA(1/2) + C prot
Immunization Groups

gp140.modSF162.delV2

gaattcgccaccatggatgcaatgaagagagggctctgctgtgtgctgctgctgtgtggagcagtc
ttcgtttcgcccagcgccgtggagaagctgtgggtgaccgtgtactacggcgtgcccgtgtggaag
gaggccaccaccaccctgttctgcgccagcgacgccaaggcctacgacaccgaggtgcacaacgtg
tgggccacccacgcctgcgtgcccaccgaccccaacccccaggagatcgtgctggagaacgtgacc
gagaacttcaacatgtggaagaacaacatggtggagcagatgcacgaggacatcatcagcctgtgg
gaccagagcctgaagccctgcgtgaagctgaccccccctgtgcgtgaccctgcactgcaccaacctg
aagaacgccaccaacaccaagagcagcaactggaaggagatggagccgcggcgagatcaagaactgc
agcttcaaggtgggcgccggcaagctgatcaactgcaacaccagcgtgatcacccaggcctgcccc
aaggtgagcttcgagcccatccccatccactactgcgccccccgccggcttcgccatcctgaagtgc
aacgacaagaagttcaacggcagcggccccctgcaccaacgtgagcaccgtgcagtgcacccacggc
atccgccccgtggtgagcacccagctgctgctgaacggcagcctggccgaggagggcgtggtgatc
cgcagcgagaacttcaccgacaacgccaagaccatcatcgtgcagctgaaggagagcgtggagatc
aactgcacccgccccaacaacaacacccgcaagagcatcaccatcggccccggccgcgccttctac
gccaccggcgacatcatcggcgacatccgccaggcccactgcaacatcagcggcgagaagtggaac
aacaccctgaagcagatcgtgaccaagctgcaggcccagttcggcaacaagaccatcgtgttcaag
cagagcagcggcggcgacccccgagatcgtgatgcacagcttcaactgcggcggcgagttctttctac
tgcaacagcacccagctgttcaacagcacctggaacaacaccatcggcccccaacaacaccaacggc
accatcaccctgccctgccgcatcaagcagatcatcaacccgctggcaggaggtgggcaaggccatg
tacgccccccccatccgcggccagatccgctgcagcagcaacatcaccggcctgctgctgacccgc
gacggcggcaaggagatcagcaacaccaccgagatctttccgcccccggcggcggcgacatgcgcgac
aactggcgcagcgagctgtacaagtacaaggtggtgaagatcgagccccctgggcgtggccccacc
aaggccaagcgccgcgtggtgcagcgcgagaagcgcgccgtgaccctgggcgccatgttcctgggc
ttcctgggcgccggccggcagcaccatgggcgcccgcagcctgaccctgaccgtgcaggcccgccag
ctgctgagcggcatcgtgcagcagcagaacaacctgctgcgcgccatcgaggcccagcagcacctg
ctgcagctgaccgtgtggggcatcaagcagctgcaggcccgcgtgctggccgtggagcgctacctg
aaggaccagcagctgctgggcatctggggctgcagcggcaagctgatctgcaccaccgccgtgccc
tggaacgccagctggagcaacaagagcctggaccagatctggaacaacatgacctggatggagtgg
gagcgcgagatcgacaactacaccaacctgatctacacctgatcgaggagagccagaaccagcag
gagaagaacgagcaggagctgctggagctggacaagtgggccagcctgtggaactggttcgacatc
agcaagtggctgtggtacatctaactcgag

Figure 6

gp140.mut7.modSF162.delV2

gaattcgccaccatggatgcaatgaagagagggctctgctgtgtgctgctgctgtgtggagcagtc
ttcgtttcgcccagcgccgtggagaagctgtgggtgaccgtgtactacggcgtgcccgtgtggaag
gaggccaccaccaccctgttctgcgccagcgacgccaaggcctacgacaccgaggtgcacaacgtg
tgggccacccacgcctgcgtgcccaccgaccccaacccccaggagatcgtgctggagaacgtgacc
gagaacttcaacatgtggaagaacaacatggtggagcagatgcacgaggacatcatcagcctgtgg
gaccagagcctgaagccctgcgtgaagctgaccccccctgtgcgtgaccctgcactgcaccaacctg
aagaacgccaccaacaccaagagcagcaactggaaggagatggaccgcggcgagatcaagaactgc
agcttcaaggtgggcgccggcaagctgatcaactccaacaccagcgtgatcacccaggcctgcccc
aaggtgagcttcgagcccatccccatccactactgcgcccccgccggcttcgccatcctgaagtgc
aacgacaagaagttcaacggcagcggcccctgcaccaacgtgagcaccgtgcagtgcacccacggc
atccgccccgtggtgagcacccagctgctgctgaacggcagcctggccgaggagggcgtggtgatc
cgcagcgagaacttcaccgacaacgccaagaccatcatcgtgcagctgaaggagagcgtggagatc
aactgcacccgccccaacaacaacaccccgcaagagcatcaccatcggccccggccgcgccttctac
gccaccggcgacatcatcggcgacatccgccaggcccactgcaacatcagcggcgagaagtggaac
aacaccctgaagcagatcgtgaccaagctgcaggcccagttcggcaacaagaccatcgtgttcaag
cagagcagcggcggcgacccccgagatcgtgatgcacagcttcaactgcggcggcgagttcttctac
tgcaacagcacccagctgttcaacagcacctggaacaacaccatcggccccaacaacaccaacggc
accatcaccctgccctgccgcatcaagcagatcatcaaccgctggcaggagggtgggcaaggccatg
tacgccccccccatccgcggccagatccgctgcagcagcaacatcaccggcctgctgctgacccgc
gacggcggcaaggagatcagcaacaccaccgagatcttccgccccggcggcggcgacatgcgcgac
aactggcgcagcgagctgtacaagtacaaggtggtgaagatcgagcccctgggcgtggccccacc
aaggccatcagcagccgtggtgcagagcgagaagagcgccgtgaccctgggcgccatgttcctgggc
ttcctgggcgccgccggcagcaccatgggcgccccgcagcctgaccctgaccgtgcaggcccgccag
ctgctgagcggcatcgtgcagcagcagaacaacctgctgcgcgccatcgaggcccagcagcacctg
ctgcagctgaccgtgtggggcatcaagcagctgcaggcccgcgtgctggccgtggagcgctacctg
aaggaccagcagctgctgggcatctggggctgcagcggcaagctgatctgccaccaccgccgtgccc
tggaacgccagctggagcaacaagagcctggaccagatctggaacaacatgacctggatggagtgg
gagcgcgagatcgacaactacaccaacctgatctacaccctgatcgaggagagccagaaccagcag
gagaagaacgagcaggagctgctggagctggacaagtgggccagcctgtggaactggttcgacatc
agcaagtggctgtggtacatctaactcgag

Figure 7

gp140mod.TV1.delV2

```
   1 gaattcatgc gcgtgatggg cacccagaag aactgccagc agtggtggat ctggggcatc
  61 ctgggctttct ggatgctgat gatctgcaac acggaggacc tgtgggtgac cgtgtactac
 121 ggcgtgcccg tgtggcgcga cgccaagacc accctgttct gcgccagcga cgccaaggcc
 181 tacgagaccg aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac
 241 ccccaggaga tcgtgctggg caacgtgacc gagaacttca acatgtggaa gaacgacatg
 301 gccgaccaga tgcacgagga cgtgatcagc ctgtgggacc agagcctgaa gccctgcgtg
 361 aagctgaccc ccctgtgcgt gacccctgaac tgcaccgaca ccaacgtgac cggcaaccgc
 421 accgtgaccg gcaacagcac caacaacacc aacggcaccg gcatctacaa catcgaggag
 481 atgaagaact gcagcttcaa cgccggcgcc ggccgcctga tcaactgcaa caccagcacc
 541 atcacccagg cctgccccaa ggtgagcttc gaccccatcc ccatccacta ctgcgccccc
 601 gccggctacg ccatcctgaa gtgcaacaac aagaccttca acggcaccgg cccctgctac
 661 aacgtgagca ccgtgcagtg cacccacggc atcaagcccg tggtgagcac ccagctgctg
 721 ctgaacggca gcctggccga ggagggcatc atcatccgca gcgagaacct gaccgagaac
 781 accaagacca tcatcgtgca cctgaacgag agcgtggaga tcaactgcac ccgccccaac
 841 aacaacaccc gcaagagcgt gcgcatcggc cccggccagg ccttctacgc caccaacgac
 901 gtgatcggca acatccgcca ggcccactgc aacatcagca ccgaccgctg gaacaagacc
 961 ctgcagcagg tgatgaagaa gctgggcgag cacttcccca acaagaccat ccagttcaag
1021 ccccacgccg gcggcgacct ggagatcacc atgcacagct tcaactgccg cggcgagttc
1081 ttctactgca acaccagcaa cctgttcaac agcacctacc acagcaacaa cggcacctac
1141 aagtacaacg gcaacagcag cagccccatc accctgcagt gcaagatcaa gcagatcgtg
1201 cgcatgtggc agggcgtggg ccagggccac tacgccccca ccatcgccag caacatcacc
1261 tgccgcagca acatcaccgg catcctgctg acccgcgacg gcggcttcaa caccaccaac
1321 aacaccgaga ccttccgccc cggcggcggc gacatgcgcg acaactggcg cagcgagctg
1381 tacaagtaca aggtggtgga gatcaagccc ctgggcatcg cccccaccaa ggccaagcgc
1441 cgcgtggtgc agcgcgagaa gcgcgccgtg ggcatcggcg ccgtgttcct gggcttcctg
1501 ggcgccgccg gcagcaccat gggcgccgcc agcatcaccc tgaccgtgca ggcccgccag
1561 ctgctgagcg gcatcgtgca gcagcagagc aacctgctga aggccatcga ggcccagcag
1621 cacatgctgc agctgaccgt gtggggcatc aagcagctgc aggccgcgt gctggccatc
1681 gagcgctacc tgaaggacca gcagctgctg ggcatctggg gctgcagcgg ccgcctgatc
1741 tgcaccaccg ccgtgccctg gaacagcagc tggagcaaca agagcgagaa ggacatctgg
1801 gacaacatga cctggatgca gtgggaccgc gagatcagca actacaccgg cctgatctac
1861 aacctgctgg aggacagcca gaaccagcag gagaagaacg agaaggacct gctggagctg
1921 gacaagtgga acaacctgtg gaactggttc gacatcagca actggccctg gtacatctaa
1981 ctcgag
```

Figure 8

gp140mod.TV1.mut7.delV2

```
   1 gaattcatgc gcgtgatggg cacccagaag aactgccagc agtggtggat ctggggcatc
  61 ctgggcttct ggatgctgat gatctgcaac accgaggacc tgtgggtgac cgtgtactac
 121 ggcgtgcccg tgtggcgcga cgccaagacc accctgttct gcgccagcga cgccaaggcc
 181 tacgagaccg aggtgcacaa cgtgtgggcc acccacgcct gcgtgcccac cgaccccaac
 241 ccccaggaga tcgtgctggg caacgtgacc gagaacttca acatgtggaa gaacgacatg
 301 gtcgaccaga tgcacgagga cgtgatcagc ctgtgggacc agagcctgaa gcctgcgtg
 361 aagctgaccc ccctgtgcgt gacctgaac tgcaccgaca ccaacgtgac cggcaaccgc
 421 accgtgaccg gcaacagcac caacaacacc aacggcaccg gcatctacaa catcgaggag
 481 atgaagaact gcagcttcaa cgccggcgcc ggccgcctga tcaactgcaa caccagcacc
 541 atcacccagg cctgccccaa ggtgagcttc gaccccatcc ccatccacta ctgcgccccc
 601 gccggctacg ccatcctgaa gtgcaacaac aagaccttca acggcaccgg ccctgctac
 661 aacgtgagca ccgtgcagtg cacccacggc atcaagcccg tggtgagcac ccagctgctg
 721 ctgaacggca gcctggccga ggaggtcatc atcatccgca gcgagaacct gaccgagaac
 781 accaagacca tcatcgtgca cctgaacgag agcgtggaga tcaactgcac ccgccccaac
 841 aacaacaccc gcaagagcgt gcgcatcggc cccggccagg ccttctacgc caccaacgac
 901 gtgatcggca acatccgcca ggcccactgc aacatcagca ccgaccgctg gaacaagacc
 961 ctgcagcagg tgatgaagaa gctgggcgag cacttcccca acaagaccat ccagttcaag
1021 ccccacgccg gcggcgacct ggagatcacc atgcacagct tcaactgccg cggcgagttc
1081 ttctactgca acaccagcaa cctgttcaac agcacctacc acagcaacaa cggcacctac
1141 aagtacaacg gcaacagcag cagccccatc accctgcagt gcaagatcaa gcagatcgtg
1201 cgcatgtggc agggcgtggg ccaggccacc tacgcccccc ccatcgccgg caacatcacc
1261 tgccgcagca acatcaccgg catcctgctg acccgcgacg gcggcttcaa caccaccaac
1321 aacaccgaga ccttccgccc cggcggcggc gacatgcgcg acaactggcg cagcgagctg
1381 tacaagtaca aggtggtgga gatcaagccc ctgggcatcg cccccaccaa ggccatcagc
1441 agcgtggtgc agagcgagaa gagcgccgtg ggcatcggcg ccgtgttcct gggcttcctg
1501 ggcgccgccg gcagcaccat gggcgccgcc agcatcaccc tgaccgtgca ggcccgccag
1561 ctgctgagcg gcatcgtgca gcagcagagc aacctgctga aggccatcga ggcccagcag
1621 cacatgctgc agctgaccgt gtggggcatc aagcagctgc aggcccgcgt gctggccatc
1681 gagcgctacc tgaaggacca gcagctgctg ggcatctggg gctgcagcgg ccgcctgatc
1741 tgcaccaccg ccgtgccctg gaacagcagc tggagcaaca agagcgagaa ggacatctgg
1801 gacaacatga cctggatgca gtgggaccgc gagatcagca actacaccgg cctgatctac
1861 aacctgctgg aggacagcca gaaccagcag gagaagaacg agaaggacct gctggagctg
1921 gacaagtgga acaacctgtg gaactggttc gacatcagca actggccctg gtacatctaa
1981 ctcgag
```

Figure 9

FIGURE 10
gp160mod.Q23-17

```
   1 ATGCGCGTGA TGGGCATCCA GCGCAACTGC CAGCACCTGC TGACCTGGGG CATCATGATC
  61 CTGGGCACCA TCATCTTCTG CAGCGCCGTG GAGAACCTGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGCGCGACGC CGACACCACC CTGTTCTGCG CCAGCGACGC CAAGGCCTAC
 181 GAGACCGAGA AGCACAACGT GTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CAGGAGATCC ACCTGGACAA CGTGACCGAG AAGTTCAACA TGTGGAAGAA CAACATGGTG
 301 GAGCAGATGC ACACCGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGCACTGC ACCAACGTGA CCAGCGTGAA CACCACCGGC
 421 GACCGCGAGG GCCTGAAGAA CTGCAGCTTC AACATGACCA CCGAGCTGCG CGACAAGCGC
 481 CAGAAGGTGT ACAGCCTGTT CTACCGCCTG GACATCGTGC CCATCAACGA GAACCAGGGC
 541 AGCGAGTACC GCCTGATCAA CTGCAACACC AGCGCCATCA CCCAGGCCTG CCCCAAGGTG
 601 AGCTTCGAGC CCATCCCCAT CCACTACTGC ACCCCCGCCG GCTTCGCCAT CCTGAAGTGC
 661 AAGGACGAGG GCTTCAACGG CACCGGCCTG TGCAAGAACG TGAGCACCGT GCAGTGCACC
 721 CACGGCATCA AGCCCGTGGT GAGCACCCAG CTGCTGCTGA ACGGCAGCCT GGCCGAGAAG
 781 AACATCACCA TCCGCAGCGA GAACATCACC AACAACGCCA AGATCATCAT CGTGCAGCTG
 841 GTGCAGCCCG TGACCATCAA GTGCATCCGC CCCAACAACA ACACCCGCAA GAGCATCCGC
 901 ATCGGCCCCG GCCAGGCCTT CTACGCCACC GGCGACATCA TCGGCGACAT CCGCCAGGCC
 961 CACTGCAACG TGACCCGCAG CCGCTGGAAC AAGACCCTGC AGGAGGTGGC CGAGAAGCTG
1021 CGCACCTACT TCGGCAACAA GACCATCATC TTCGCCAACA GCAGCGGCGG CGACCTGGAG
1081 ATCACCACCC ACAGCTTCAA CTGCGGCGGC GAGTTCTTCT ACTGCAACAC CAGCGGCCTG
1141 TTCAACAGCA CCTGGTACGT GAACAGCACC TGGAACGACA CCGACAGCAC CCAGGAGAGC
1201 AACGACACCA TCACCCTGCC CTGCCGCATC AAGCAGATCA TCAACATGTG GCAGCGGGCC
1261 GGCCAGGCCA TGTACGCCCC CCCCATCCCC GGCGTGATCA AGTGCGAGAG CAACATCACC
1321 GGCCTGCTGC TGACCCGCGA CGGCGGCAAG GACAACAACG TGAACGAGAC CTTCCGCCCC
1381 GGCGGCGGCG ACATGCGCGA CAACTGGCGC AGCGAGCTGT ACAAGTACAA GGTGGTGGAG
1441 ATCGAGCCCC TGGGCGTGGC CCCCACCCGC GCCAAGCGCC GCGTGGTGGA GCGCGAGAAG
1501 CGCGCCGTGG GCATCGGCGC CGTGTTCCTG GGCTTCCTGG GCGCCGCCGG CAGCACCATG
1561 GGCGCCACCA GCATCACCCT GACCGTGCAG GCCCGCCAGC TGCTGAGCGG CATCGTGCAG
1621 CAGCAGAACA ACCTGCTGCG CGCCATCGAG GCCCAGCAGC ACCTGCTGAA GCTGACCGTG
1681 TGGGGCATCA AGCAGCTGCA GGCCCGCGTG CTGGCCGTGG AGCGCTACCT GCGCGACCAG
1741 CAGCTGCTGG GCATCTGGGG CTGCAGCGGC AAGCTGATCT GCACCACCAA CGTGCCCTGG
1801 AACAGCAGCT GGAGCAACAA GAGCCTGGAC GAGATCTGGA ACAACATGAC CTGGCTGCAG
1861 TGGGACAAGG AGATCAACAA CTACACCCAG CTGATCTACC GCCTGATCGA GGAGAGCCAG
1921 AACCAGCAGG AGAAGAACGA GAAGGAGCTG CTGGAGCTGG ACAAGTGGGC CAACCTGTGG
1981 AGCTGGTTCG ACATCAGCAA CTGGCTGTGG TACATCAAGA TCTTCATCAT CATCGTGGGC
2041 GGCCTGATCG GCCTGCGCAT CGTGTTCGCC GTGCTGAGCG TGATCAACCG CGTGCGCCAG
2101 GGCTACAGCC CCCTGAGCTT CCAGACCCAC ACCCCCAACC CCCGCGGCCT GGACCGCCCC
2161 GAGCGCATCG AGGAGGAGGA CGGCGAGCAG GGCCGCGGCC GCAGCATCCG CCTGGTGAGC
2221 GGCTTCCTGG CCCTGGCCTG GGACGACCTG CGCAGCCTGT GCCTGTTCAG CTACCACCGC
2281 CTGCGCGACT TCATCCTGAT CGCCGCCCGC ACCGTGGAGC TGCTGGGCCA CAGCAGCCTG
2341 AAGGGCCTGC GCCTGGGCTG GGAGGGCATC AAGTACCTGT GGAACCTGCT GAGCTACTGG
2401 GGCCGCGAGC TGAAGATCAG CGCCATCAAC CTGGTGGACA CCATCGCCAT CGCCGTGGCC
2461 GGCTGGACCG ACCGCGTGAT CGAGATCGCC CAGCGCATCG GCCGCGCCAT CCTGCACATC
2521 CCCGTGCGCA TCCGCCAGGG CCTGGAGCGC GCCCTGCTGT AA
```

FIGURE 11

gp160mod.98UA0116

```
   1 ATGAAGGCCC GCGGCATGCA GCGCAACTAC CAGCACCTGT GGCGCTGGGG CACCATGCTG
  61 TTCTGGATGA TCATCATGTG CAAGGCCGCC GAGAACCTGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGCGCGACGC CGAGCACCAC CTGTTCTGCG CCAGCGACGC CAAGGCCTAC
 181 GACAAGGAGG TGCACAACGT GTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCGACCCC
 241 CAGGAGATCA TCCTGGAGAA CGTGACCGAG AAGTTCAACA TGTGGAAGAA CAACATGGTG
 301 GAGCAGATGC AGACCGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGAACTGC GCCGGCCCCA GCAGCAACAA CAGCAACGTG
 421 AACAGCAACA GCAACGACAA CTGGAGCGAG GAGATGAAGA ACTGCAGCTT CAACATGACC
 481 ACCGAGCTGC GCGACAAGCG CAAGACCGTG CACAGCCTGT TCTACAAGCT GGACATCGTG
 541 AGCACCGGCA GCAACGACAG CCGCCAGTAC CGCCTGATCA ACTGCAACAC CAGCGCCATG
 601 ACCCAGGCCT GCCCCAAGGT GACCTTCGAG CCCATCCCCA TCCACTACTG CGCCCCCGCC
 661 GGCTTCGCCA TCCTGAAGTG CAAGGACACC AACTTCACCG GCACCGGCCC CTGCAAGGAC
 721 GTGAGCACCG TGCAGTGCAC CCACGGCACC AAGCCCGTGG TGAGCACCCA GCTGCTGCTG
 781 AACGGCAGCC TGGCCGAGAA GGAGGTGATG ATCCGCAGCG AGAACATCAC CGACAACGGC
 841 AAGATCATCA TCGTGCAGCT GACCGAGCCC GTGAACATCA CCCGCATCCG CCCCGGCGAG
 901 AACAAGCGCA CCAGCATCCG CATCGGCCCC GGCCAGACCT TCTACGCCAC CGGCGACGTG
 961 ATCGGCGACA TCCGCAAGGC CTACTGCAAC GTGAGCCGCG CCGCCTGGAA CAGCACCCTG
1021 CAGAAGATCA GCACCCAGCT GCGCCAGTAC TTCAACAACA AGACCATCAT CTTCAAGAAC
1081 AGCAGCGGCG GCGACCTGGA GGTGACCACC CACAGCTTCA ACTGCGGCGG CGAGTTCTTC
1141 TACTGCAACA CCACCGACCT GTTCAACAGC ACCTGGAACG AGCACGGCCC CGTGACCAAC
1201 AGCACCATGG CCAACGGCAC CATCACCCTG CCCTGCCGCA TCAAGCAGAT CATCAACATG
1261 TGGCAGCGCG TGGGCCAGGC CATGTACGCC CCCCCCATCG AGGGCAACAT CCGCTGCGAG
1321 AGCAACATCA CCGGCCTGCT GCTGACCCGC GACGGCGGCA GCGGCGCCAA CAGCAGCAAG
1381 GAGACCTACC GCCCCATCGG CGGCAACATG CGCGACAACT GGCGCAGCGA GCTGTACAAG
1441 TACAAGGTGG TGAAGATCGA GCCCATCGGC GTGGCCCCCA CCAAGGCCAA GCGCCGCGTG
1501 GTGGAGCGCG AGAAGCGCGC CATCGGCCTG GGCGCCGCCT TCCTGGGCTT CCTGGGCGCC
1561 GCCGGCAGCA CCATGGGCGC CGCCAGCATG ACCCTGACCG TGCAGGCCCG CCAGCTGCTG
1621 AGCGGCATCG TGCAGCAGCA GAGCAACCTG CTGCGCGCCA TCGAGGCCCA GCAGCACCTG
1681 CTGAAGCTGA CCGTGTGGGG CATCAAGCAG CTGCAGGCCC GCGTGCTGGC CGTGGAGCGC
1741 TACCTGAAGG ACCAGCAGCT GCTGGGCATC TGGGGCTGCA GCGGCAAGCT GATCTGCACC
1801 ACCAACGTGC CCTGGAACAG CAGCTGGAGC AACAAGAGCC AGAGCGAGAT CTGGGGCAAC
1861 ATGACCTGGA TGCAGTGGGA CCGCGAGGTG ATCAACTACA CCAACATCAT CTACGACCTG
1921 ATCGAGGAGA GCCAGAACCA GCAGGAGAAG AACGAGCAGG ACCTGCTGGC CCTGGACAAG
1981 TGGGCCAGCC TGTGGAGCTG GTTCGACATC AGCAACTGGC TGTGGTACAT CAAGATCTTC
2041 ATCATCATCG TGGGCGGCCT GATCGGCCTG CGCATCGTGT TCGCCGTGCT GAGCATCATC
2101 AACCGCGCCG GCCAGGGCTA CAGCCCCCTG AGCCTGCAGA CCCTGACCCC CCACCCCGAG
2161 GGCCCCGACC GCCCCGGCCG CATCAAGGAG GAGGGCGGCG AGCAGGACCG CGACCGCAGC
2221 ATCCGCCTGG TGAGCGGCTT CCTGGCCCTG GCCTGGGACG ACCTGCGCAG CCTGTGCCTG
2281 TTCAGCTACC GCCGCCTGCG CGACTTCATC AGCATCGCCG CCCGCACCGT GGAGCTGCTG
2341 GGCCGCAGCA GCCTGAAGGG CCTGCGCCTG GGCTGGGAGG GCCTGAAGTA CCTGGGCAAC
2401 CTGCTGGGCT ACCGCGGCCA GGAGCTGAAG AGCAGCGCCA TCAACCTGAT CGACACCATC
2461 GCCATCGCCG TGGCCGGCTG GACCGACCGC GTGATCGAGA TCGGCCAGCG CTTCTGCCGC
2521 GCCATCCGCA ACATCCCCCG CCGCATCCGC CAGGGCGCCG AGCGCGCCCT GCAGTAA
```

FIGURE 12

gp160mod.SE8538

```
   1 ATGCGCGTGA AGGGCATCCA GCGCAACAGC CAGCACCTGC TGCGCTGGGG CACCATGATC
  61 CTGGGCATGA TCATCATCTG CAGCACCGCC GACAAGCTGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGAAGGACGC CGAGACCACC CTGTTCTGCG CCAGCGACGC CAAGGCCTAC
 181 GACACCGAGG TGCACAACGT GTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CAGGAGCTGC ACCTGGCCAA CGTGACCGAG GAGTTCAACA TGTGGAAGAA CAGCATGGTG
 301 GAGCAGATGC ACACCGACAT CATCAGCCTG TGGGACCAGA GCCTGATCCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGGAGTGC AACGACTACA ACTACAACGT GACCAACAGC
 421 AGCCACAGCT ACAACGTGAC CAACATGCAG GAGATGAAGA ACTGCAGCTT CAACGTGACC
 481 ACCGAGCTGC GCGACAAGCG CCAGAAGGTG ACCAGCCTGT TCTACAAGCT GGACGTGGTG
 541 CCCATCGGCG GCAACGACAC CAACAGCACC CAGTACCGCC TGATCAACTG CAACACCAGC
 601 GCCATCACCC AGGCCTGCCC CAAGGTGACC TTCGAGCCCA TCCCCATCCA CTACTGCGCC
 661 CCCGCCGGCT TCGCCATCCT GAAGTGCCGC GACGAGAACT TCAACGGCAC CGGCCCCTGC
 721 AAGAACGTGA GCACCGTGCA GTGCACCCAC GGCATCAAGC CCGTGGTGAG CACCCAGCTG
 781 CTGCTGAACG GCAGCCTGGC CCGCGAGAAG GTGATGATCC GCAGCGAGAA CATCACCAAC
 841 AACGTGAAGA ACATCATCGT GCAGCTGAAG GAGCCCGTGG AGATCAACTG CACCCGCCCC
 901 GGCAACAACA CCCGCAAGAG CATCCGCATC GGCCCCGGCC AGGCCTTCTA CGCCACCGGC
 961 GAGGTGATCG GCGACATCCG CCAGGCCCAC TGCAACGTGA GCCGCGCCAA GTGGAACAAG
1021 ACCCTGCACG AGGTGGCCAA GCAGCTGCGC ACCTACTTCA ACAACAAGAC CATCATCTTC
1081 ACCAACAGCA GCGGCGGCGA CCTGGAGATC ACCACCCACA CCGTGAACTG CGGCGGCGAG
1141 TTCTTCTACT GCAACACCAG CGGCCTGTTC AACAGCACCT GGAGCAGCAA CGCCAGCGAG
1201 CCCATGAGCA ACAGCACCGA GAGCAACGAC ACCATCACCC TGCAGTGCCG CATCCGCCAG
1261 ATCATCAACA TGTGGCAGCG CGCCGGCAAG GCCATCTACG CCCCCCCCAT CCCCGGCATC
1321 ATCAAGTGCG TGAGCAACAT CACCGGCCTG ATCCTGACCC GCGACGGCGG CAGCAACAAC
1381 AGCACCAACG AGACCTTCCG CCCCGGCGGC GGCGACATGC GCGACAACTG GCGCAGCGAG
1441 CTGTACAAGT ACAAGGTGGT GAAGATCGAG CCCCTGGGCG TGGCCCCCAC CAAGGCCAAG
1501 CGCCGCGTGG TGGAGCGCGA GAAGCGCGCC ATCGGCATCG GCGCCGTGTT CATCGGCTTC
1561 CTGGGCGCCG CCGGCAGCAC CATGGGCGCC GCCAGCATCA CCCTGACCGT GCAGGCCCGC
1621 CAGCTGCTGA GCGGCATCGT GCAGCAGCAG AGCAACCTGC TGCGCGCCAT CGAGGCCCAG
1681 CAGCACCTGC TGAAGCTGAC CGTGTGGGGC ATCAAGCAGC TGCAGGCCCG CGTGCTGGCC
1741 GTGGAGCGCT ACCTGAAGGA CCAGCAGCTG CTGGGCATCT GGGGCTGCAG CGGCAAGCTG
1801 ATCTGCACCA CCAACGTGCC CTGGAACAGC AGCTGGAGCA ACAAGAGCCA GAGCGAGATC
1861 TGGGACAACA TGACCTGGCT GCAGTGGGAC AAGGAGATCA GCAACTACAC CCAGACCATC
1921 TACCGCCTGA TCGAGGAGAG CCAGAACCAG CAGGAGAAGA ACGAGCAGGA CCTGCTGGCC
1981 CTGGACAAGT GGGCCAGCCT GTGGAACTGG TTCGACATCA GCCGCTGGCT GTGGTACATC
2041 CGCATCTTCA TCATGATCGT GGGCGGCCTG ATCGGCCTGC GCATCGTGTT CGCCGTGCTG
2101 AGCGTGATCA ACCGCGTGCG CCAGGGCTAC AGCCCCCTGA GCTTCCAGAT CCACACCCCC
2161 AACCCCGGCG ACCTGGACCG CCCCGGCCGC ATCGAGGAGG AGGGCGGCGA GCAGGACCGC
2221 GGCCGCAGCA TCCGCCTGGT GAGCGGCTTC CTGGCCCTGG CCTGGGACGA CCTGCGCAGC
2281 CTGTGCCTGT TCAGCTACCA CCGCCTGCGC GACTTCATCC TGATCGCCGC CCGCACCGTG
2341 GAGCTGCTGG GCCAGCGCGG CTGGGAGGGC CTGAAGTACC TGTGGAACCT GCTGGTGTAC
2401 TGGATCCGCG AGCTGAAGAT CAGCGCCATC AGCCTGCTGG ACACCATCGC CATCGCCGTG
2461 GCCGGCTGGA CCGACCGCGT GATCGAGCTG GGCCAGCGCC TGTGCCGCGC CATCCTGCAC
2521 ATCCCCGTGC GCATCCGCCA GGGCTTCGAG CGCGCCCTGC TGTAA
```

FIGURE 13

gp160mod.UG031

```
   1 ATGCGCGTGC GCGGCATCCA GACCAGCTGG CAGAACCTGT GGCGCTGGGG CACCATGATC
  61 CTGGGCATGC TGATGATCTA CAGCGCCGCC GAGAACCTGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGAAGGACGC CGAGACCACC CTGTTCTGCG CCAGCGACGC CAAGGCCTAC
 181 GACACCGAGG TGCACAACGT GTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CAGGAGATCC ACCTGGAGAA CGTGACCGAG GACTTCAACA TGTGGAAGAA CAACATGGTG
 301 GAGCAGATGC ACACCGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGGAG
 361 CTGACCCCCC TGTGCGTGAC CCTGGACTGC CTGAACGCCA CCCTGAACGC CACCGCCCCC
 421 AACGTGACCA ACGACATGGA GGGCGAGATG AAGAACTGCA GCTACAACAT CACCACCGAG
 481 CTGAAGGACA AGAAGCAGCA GGTGTACAGC CTGTTCTACA AGCTGGACGT GGTGCAGATC
 541 AACGAGAAGA ACAAGACCAA CAAGTACCGC CTGATCAACT GCAACACCAG CGCCATCACC
 601 CAGGCCTGCC CCAAGGTGAG CTTCGAGCCC ATCCCCATCC ACTACTGCGC CCCCGCCGGC
 661 TTCGCCATCC TGAAGTGCAA GGACACCGAG TTCAACGGCA CCGGCCCCTG CAAGAACGTG
 721 AGCACCGTGC AGTGCACCCA CGGCATCCGC CCCGTGATCA GCACCCAGCT GCTGCTGAAC
 781 GGCAGCCTGG CCGAGGGCGG CATCCAGATC CGCAGCGAGA ACATCACCAA CAACGCCAAG
 841 ACCATCATCG TGCAGCTGGA CAAGGCCGTG AAGATCAACT GCACCCGCCC CAACAACAAC
 901 ACCCGCAAGA GCGTGCGCAT CGGCCCCGGC CAGGCCTTCT ACGCCACCGG CGACATCATC
 961 GGCGACATCC GCCAGGCCCA CTGCAACGTG AGCCGCGCCA AGTGGAACGA GACCCTGCGC
1021 GGCATCGCCA AGAAGCTGAG CGAGCACTTC AAGAACAAGA TCATCATCTT CGAGAAGAGC
1081 AGCGGCGGCG ACATCGAGAT CACCACCCAC AGCTTCAACT GCGGCGGCGA GTTCTTCTAC
1141 TGCAACACCA GCGGCCTGTT CAACGGCACC TGGAAGCCCA ACAGCACCGA GAGCAACAAC
1201 ACCACCCCCA ACGACACCAT CACCCTGACC TGCCGCATCA AGCAGATCAT CAACATGTGG
1261 CAGAAGGTGG GCCAGGCCAT GTACGCCCCC CCCATCCAGG GCGTGATCCG CTGCGAGAGC
1321 AACATCACCG GCCTGCTGCT GACCCGCGAC GGCGGCATCA ACAGCATCAA CGAGACCTTC
1381 CGCCCCGGCG GCGGCAACAT GCGCGACAAC TGGCGCAGCG AGCTGTACAA GTACAAGGTG
1441 GTGAAGATCG AGCCCCTGGG CGTGGCCCCC AGCCGCGCCA AGCGCCGCGT GGTGGAGCGC
1501 GAGAAGCGCG CCGTGGGCAT CGGCGCCGTG TTCCTGGGCT TCCTGGGCGC CGCCGGCAGC
1561 ACCATGGGCG CCGCCAGCAT CACCCTGACC GCCCAGGCCC GCCAGCTGCT GAGCGGCATC
1621 GTGCAGCAGC AGAGCAACCT GCTGCGCGCC ATCAAGGCCC AGCAGCACAT GCTGAAGCTG
1681 ACCGTGTGGG GCATCAAGCA GCTGCAGGCC CGCGTGCTGG CCGTGGAGCG CTACCTGAAG
1741 GACCAGCAGC TGCTGGGCAT CTGGGGCTGC AGCGGCAAGC TGATCTGCAC CACCAACGTG
1801 CCCTGGAACA GCAGCTGGAG CAACAAGAGC ATGAACGAGA TCTGGGACAA CATGACCTGG
1861 CTGCAGTGGG AGAAGGAGAT CAGCAACTAC ACCCAGCTGA TCTACAACCT GATCGAGGAG
1921 AGCCAGAACC AGCAGGAGAA GAACGAGCAG GACCTGCTGG CCCTGGACAA GTGGGCCAGC
1981 CTGTGGAACT GGTTCGACAT CAGCCGCTGG CTGTGGTACA TCAAGATCTT CATCATGATC
2041 GTGGGCGGCC TGATCGGCCT GCGCATCGTG TTCGCCGTGC TGAGCGTGAT CAACCGCGTG
2101 CGCCAGGGCT ACAGCCCCCT GAGCTTCCAG ATCCGCACCC CCAACCCCGA GGAGCCCGAC
2161 CGCCTGGGCC GCATCGGCGA GGAGGGCGGC GAGCAGGACC GCGACCGCAG CATCCGCCTG
2221 GTGAGCGGCT TCCTGGCCCT GGCCTGGGAC GACCTGCGCA GCCTGTGCCT GTTCAGCTAC
2281 CACCGCCTGC GCGACTTCAT CAGCATCGCC GCCCGCACCG TGGAGCTGCT GGGCCACAGC
2341 AGCCTGAAGG GCCTGCGCCT GGGCTGGGAG GGCCTGAAGT ACCTGTGGAA CCTGCTGCTG
2401 TACTGGGGCC TGGAGCTGAA GACCAGCGCC GTGAACCTGG TGGACACCAT CGCCATCGCC
2461 GTGGCCGGCT GGACCGACCG CGTGATCGAG ATCGGCCAGC GCATCTTCCG CGCCATCCTG
2521 AACATCCCCC GCCGCATCCG CCAGGGCCTG GAGCGCGGCC TGCTGTAA
```

FIGURE 14

gp160mod.92UG001

```
   1 ATGCGCGTGC GCGAGATCGA GCGCAACTAC CTGTGCCTGT GGCGCTGGGG CATCATGCTG
  61 CTGGGCATGC TGATGACCTA CAGCGTGGCC GAGAAGAAGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGAAGGAGGC CACCACCACC CTGTTCTGCG CCAGCGACGC CAAGAGCTAC
 181 AAGACCGAGG TGCACAACAT CTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CGCGAGATCG AGCTGGAGAA CGTGACCGAG AACTTCAACA TGTGGAAGAA CAACATGGTG
 301 GAGCAGATGC ACGAGGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGAACTGC ACCGACGCCC GCCGCAACGA GACCCGCAAC
 421 AACATCACCG GCATGGAGAA CAACGACCAG ATCGAGATGA AGAACTGCAG CTTCAACATC
 481 ACCACCAAGC TGATCGACAA GAAGAAGCAG GTGCACGCCC TGTTCTACCG CCTGGACGTG
 541 GTGCAGATCG ACAACGACAC CAGCAACAGC AACTACAGCA ACTACCGCCT GATCAACTGC
 601 AACACCAGCG CCATCACCCA GGCCTGCCCC AAGGTGACCT TCGAGCCCAT CCCCATCCAC
 661 TACTGCGCCC CCGCCGGCTT CGCCATCCTG AAGTGCCGCG ACAAGAAGTT CAACGGCACC
 721 GGCCCCTGCA AGAACGTGAG CACCGTGCAG TGCACCCACG GCATCCGCCC CGTGGTGAGC
 781 ACCCAGCTGC TGCTGAACGG CAGCCTGGCC GAGGAGGAGA TCATCATCCG CAGCGAGAAC
 841 CTGACCAACA ACGCCAAGAC CCTGATCGTG CAGCTGAACG AGAGCGTGGA GATCAACTGC
 901 ACCCGCCCCT ACTACAACCA GATCCGCCAG CGCACCAGCA TCGGCCAGGG CCAGGCCCTG
 961 TACACCACCC GCGTGACCGG CGACATCCGC AAGGCCTACT GCAACATCAG CAAGGCCGGC
1021 TGGAACAAGA CCCTGCAGCA GGTGGCCAAG AAGCTGGGCG ACCTGTTCAA CCAGACCACC
1081 ATCATCTTCA AGCCCAGCAG CGGCGGCGAC CCCGAGATCA CCACCCACAG CTTCAACTGC
1141 GGCGGCGAGT TCTTCTACTG CAACACCAGC AAGCTGTTCA ACAGCGCCTG GAACGACAGC
1201 ACCTGGAACA TCGGCAACAA CAACACCGGC AGCGACAACG AGACCATCAT CATCCCCTGC
1261 CGCATCAAGC AGATCATCAA CATGTGGCAG GGCGTGGGCA AGGCCATGTA CGCCCCCCCC
1321 ATCGAGGGCT GGATCAACTG CGCCAGCAAC ATCACCGGCC TGCTGCTGGT GCGCGACGGC
1381 GGCGGCGCCA ACGACAGCCA GAACGAGACC TTCCGCCCCC AGGGCGGCGA CATGCGGGAC
1441 AACTGGCGCA GCGAGCTGTA CAAGTACAAG GTGGTGAAGA TCGAGCCCCT GGGCATCGCC
1501 CCCACCAAGG CCAAGCGCCG CGTGGTGGAG CGCGAGAAGC GCGCCATCGG CCTGGGCGCC
1561 ATGTTCCTGG GCTTCCTGGG CGCCGCCGGC AGCACCATGG GCGCCGCCAG CCTGACCCTG
1621 ACCGTGCAGG CCCGCCAGCT GCTGAGCGGC ATCGTGCAGC ACCAGAACAA CCTGCTGATG
1681 GCCATCGAGG CCCAGCAGCA CCTGCTGCAG CTGACCGTGT GGGGCATCAA GCAGCTGCAG
1741 GCCCGCATCC TGGCCGTGGA GCGCTACCTG CAGGACCAGC AGCTGCTGGG CAGCTGGGGC
1801 TGCAGCGGCC GCCACATCTG CACCACCACC GTGCCCTGGA ACAGCAGCTG GAGCAACAAG
1861 AGCATCGACG ACATCTGGAA CAACATGACC TGGATGGAGT GGGAGAAGGA GATCGACAAC
1921 TACACCGGCG TGATCTACCG CCTGATCGAG GAGAGCCAGA CCCAGCAGGA GAAGAACGAG
1981 CAGGAGCTGC TGCAGCTGGA CAAGTGGGCC AGCCTGTGGA ACTGGTTCAG CATCACCAAG
2041 TGGCTGTGGT ACATCAAGAT CTTCATCATG ATCGTGGGCG GCCTGATCGG CCTGCGCATC
2101 GTGTTCACCG TGCTGAGCCT GGTGAACCGC GTGCGCCAGG GCTACAGCCC CCTGAGCTTC
2161 CAGACCCTGT TCCCCGCCCC CCGCGGCCCC GACCGCCCCG AGGAGATCGA GGAGGGCGGC
2221 GGCGAGCAGG GCCGCGGCCG CAGCACCCGC CTGGTGAACG GCTTCAGCAC CCTGATCTGG
2281 GACGACCTGC GCAACCTGTG CCTGTTCAGC TACCACCGCC TGCGCGACCT GATCCTGATC
2341 GCCACCCGCA TCGTGGAGCT GCTGGGCCGC CGCGGCTGGG AGGCCATCAA GTACCTGTGG
2401 AACCTGCTGC AGTACTGGAG CCAGGAGCTG AAGACCAGCG CCATCAGCCT GTTCAACGCC
2461 ACCGCCGTGG CCGTGGCCGA GGGCACCGAC CGCGTGATCG AGGTGGTGCA GCGCTTCTTC
2521 CGCGGCATCC TGAACGTGCC CACCCGCATC CGCCAGGGCC TGGAGCGCGC CCTGCTGTAA
```

FIGURE 15

gp160mod.94UG114

```
   1 ATGCGCGTGC GCGAGACCAA GCGCAACTAC CAGCACCTGT GGAAGTGGGG CACCATGCTG
  61 CTGGGCATGC TGATGATCTG CAGCGTGACC GGCAAGAGCT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGAAGGAGGC CACCACCACC CTGTTCTGCG CCAGCGACGC CAAGGCCTAC
 181 AAGGCCGAGG CCCACAACAT CTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CAGGAGATCA AGCTGGAGAA CGTGACCGAG AACTTCAACA TGTGGAAGAA CAACATGGTG
 301 GAGCAGATGC ACGAGGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGAACTGC ACCAACTGGG TGACCGACAC CACCAACACC
 421 ACCGGCATGG CCAACTGCAG CTTCAACATC ACCACCGAGA TCCGCGACAA GAAGAAGCAG
 481 GTGCAGGCCC TGTTCTACAA GCTGGACGTG GTGAAGATCA ACGACAACGA CAGCGACAAC
 541 ACCAGCTACC GCCTGATCAA CTGCAACACC AGCGCCATCA CCCAGGCCTG CCCCAAGATG
 601 ACCTTCGAGC CCATCCCCAT CCACTACTGC GCCCCCGCCG GCTTCGCCAT CCTGAAGTGC
 661 AACGAGAAGA AGTTCAACGG CACCGGCCCC TGCAAGAACG TGAGCACCGT GCAGTGCACC
 721 CACGGCATCA AGCCCGTGGT GAGCACCCAG CTGCTGCTGA ACGGCAGCCT GGCCGAGGAG
 781 GAGATCATCA TCCGCAGCGA GAACCTGACC AACAACGCCA AGATCATCAT CGTGCAGCTG
 841 AACGAGAGCG TGCCCATCAA CTGCATCCGC CCCTACAACA ACACCCGCCA GAGCACCCGC
 901 ATCGGCCCCG GCCAGGCCCT GTTCACCACC AAGGTGATCG GCGACATCCG CCAGGCCCAC
 961 TGCAACATCA GCGGCGCCGG CTGGAACAAG ACCCTGCAGC AGGTGGCCGA GAAGCTGGGC
1021 AACCTGCTGA ACCAGACCAC CATCATCTTC AAGCCCAGCA GCGGCGGCGA CCCCGAGATC
1081 ACCACCCACA GCTTCAACTG CGGCGGCGAG TTCTTCTACT GCAACACCAC CCGCCTGTTC
1141 AACAGCACCT GGAAGCGCAA CAACAGCGAG TGGCGCAGCG ACAACACCCC CGACGAGACC
1201 ATCACCCTGC AGTGCCGCAT CAAGCAGATC ATCAACATGT GGCAGGAGGT GGGCAAGGCC
1261 ATGTACGCCC CCCCCATCGA GGGCTTCATC AACTGCAGCA GCAACATCAC CGGCCTGCTG
1321 CTGACCCGCG ACGGCGGCGC CATCAACAGC AGCCAGAACG AGACCTTCCG CCCCGGCGGC
1381 GGCGACATGC GCAACAACTG GCGCAGCGAG CTGTACAAGT ACAAGGTGGT GAAGCTGGAG
1441 CCCATCGGCC TGGCCCCCAC CGCCGCCAAG CGCCGCGTGG TGGAGGCGGA GAAGCGCGCC
1501 ATCGGCCTGG GCGCCCTGTT CCTGGGCTTC CTGGGCACCG CCGGCAGCAC CATGGGCGCC
1561 GTGAGCCTGA CCCTGACCGT GCAGGCCCGC CAGGTGCTGA GCGGCATCGT GCAGCAGCAG
1621 AACAACCTGC TGCGCGCCAT CGAGGCCCAG CAGCACCTGC TGCAGCTGAC CGTGTGGGGC
1681 ATCAAGCAGC TGCAGGCCCG CATCCTGGCC GTGGAGAGCT ACCTGAAGGA CCAGCAGCTG
1741 CTGGGCATCT GGGGCTGCAG CGGCAAGCAC ATCTGCACCA CCAACGTGCC CTGGAACAGC
1801 AGCTGGAGCA ACCGCAGCGT GGACGAGATC TGGAACAACA TGACCTGGAT GGAGTGGGAG
1861 CGCGAGATCG ACAACTACAC CGAGCTGGTG TACAGCCTGC TGGAGGTGAG CCAGATCCAG
1921 CAGGAGAAGA ACGAGCAGGA GCTGCTGAAG CTGGACACCT GGGCCAGCCT GTGGAACTGG
1981 TTCAGCATCA CCCAGTGGCT GTGGTACATC AAGATCTTCA TCATGATCGT GGGCGGCCTG
2041 ATCGGCCTGC GCATCGTGTT CGCCGTGCTG AGCGTGGTGA ACCGCGTGCG CCAGGGCTAC
2101 AGCCCCCTGA GCTTCCAGAC CCTGCTGCCC GCCCCCCGCG AGCCCGACCG CCCCGAGGGC
2161 ATCGAGGAGG AGGGCGGCGA GCGCGACCGC GGCCGCAGCA TCCGCCTGGT GAACGGCCTG
2221 AGCGCCCTGA TCTGGGACGA CCTGCGCAAC CTGTGCCTGT TCAGCTACCA CCGCCTGCGC
2281 GACCTGATCC TGATCGCCGC CCGCATCGTG GAGCTGCTGG GCCGCCGCGG CTGGGAGGCC
2341 ATCAAGTACC TGTGGAACCT GCTGCAGTAC TGGATCCAGG AGCTGAAGAA CAGCGCCGTG
2401 AGCCTGTTCA ACACCATCGC CATCGCCGTG GCCGAGGGCA CCGACCGCGC CATCGAGCTG
2461 GTGCAGCGCG CCGTGCGCGC CATCCTGAAC ATCCCCGTGC GCATCCGCCA GGGCCTGGAG
2521 CGCGCCCTGC TGTAA
```

FIGURE 16
gp160mod.ELI

```
   1 ATGCGCGCCC GCGGCATCGA GCGCAACTGC CAGAACTGGT GGAAGTGGGG CATCATGCTG
  61 CTGGGCATCC TGATGACCTG CAGCGCCGCC GACAACCTGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGAAGGAGGC CACCACCACC CTGTTCTGCG CCAGCGACGC CAAGAGCTAC
 181 GAGACCGAGG CCCACAACAT CTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CAGGAGATCG CCCTGGAGAA CGTGACCGAG AACTTCAACA TGTGGAAGAA CAACATGGTG
 301 GAGCAGATGC ACGAGGACAT CATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGAACTGC AGCGACGAGC TGCGCAACAA CGGCACCATG
 421 GGCAACAACG TGACCACCGA GGAGAAGGGC ATGAAGAACT GCAGCTTCAA CGTGACCACC
 481 GTGCTGAAGG ACAAGAAGCA GCAGGTGTAC GCCCTGTTCT ACCGCCTGGA CATCGTGCCC
 541 ATCGACAACG ACAGCAGCAC CAACAGCACC AACTACCGCC TGATCAACTG CAACACCAGC
 601 GCCATCACCC AGGCCTGCCC CAAGGTGAGC TTCGAGCCCA TCCCCATCCA CTACTGCGCC
 661 CCCGCCGGCT TCGCCATCCT GAAGTGCCGC GACAAGAAGT TCAACGGCAC CGGCCCCTGC
 721 ACCAACGTGA GCACCGTGCA GTGCACCCAC GGCATCCGCC CCGTGGTGAG CACCCAGCTG
 781 CTGCTGAACG GCAGCCTGGC CGAGGAGGAG GTGATCATCC GCAGCGAGAA CCTGACCAAC
 841 AACGCCAAGA ACATCATCGC CCACCTGAAC GAGAGCGTGA AGATCACCTG CGCCGCCCC
 901 TACCAGAACA CCCGCCAGCG CACCCCCATC GGCCTGGGCC AGAGCCTGTA CACCACCCGC
 961 AGCCGCAGCA TCATCGGCCA GGCCCACTGC AACATCAGCC GCGCCCAGTG GAGCAAGACC
1021 CTGCAGCAGG TGGCCCGCAA GCTGGGCACC CTGCTGAACA AGACCATCAT CAAGTTCAAG
1081 CCCAGCAGCG GCGGCGACCC CGAGATCACC ACCCACAGCT TCAACTGCGG CGGCGAGTTC
1141 TTCTACTGCA ACACCAGCGG CCTGTTCAAC AGCACCTGGA ACATCAGCGC CTGGAACAAC
1201 ATCACCGAGA GCAACAACAG CACCAACACC AACATCACCC TGCAGTGCCG CATCAAGCAG
1261 ATCATCAACA TGGTGGCCGG CCGCAAGGCC ATCTACGCCC CCCCCATCGA GCGCAACATC
1321 CTGTGCAGCA GCAACATCAC CGGCCTGCTG CTGACCCGCG ACGGCGGCAT CAACAACAGC
1381 ACCAACGAGA CCTTCCGCCC CGGCGGCGGC GACATGCGCG ACAACTGGCG CAGCGAGCTG
1441 TACAAGTACA AGGTGGTGCA GATCGAGCCC CTGGGCGTGG CCCCCACCCG CGCCAAGCGC
1501 CGCGTGGTGG AGCGCGAGAA GCGCGCCATC GGCCTGGGCG CCATGTTCCT GGGCTTCCTG
1561 GGCGCCGCCG GCAGCACCAT GGGCGCCCGC AGCGTGACCC TGACCGTGCA GGCCCGCCAG
1621 CTGATGAGCG GCATCGTGCA GCAGCAGAAC AACCTGCTGC GCGCCATCGA GGCCCAGCAG
1681 CACCTGCTGC AGCTGACCGT GTGGGGCATC AAGCAGCTGC AGGCCCGCAT CCTGGCCGTG
1741 GAGCGCTACC TGAAGGACCA GCAGCTGCTG GGCATCTGGG GCTGCAGCGG CAAGCACATC
1801 TGCACCACCA ACGTGCCCTG GAACAGCAGC TGGAGCAACC GCAGCCTGAA CGAGATCTGG
1861 CAGAACATGA CCTGGATGGA GTGGGAGCGC GAGATCGACA ACTACACCGG CCTGATCTAC
1921 AGCCTGATCG AGGAGAGCCA GACCCAGCAG GAGAAGAACG AGAAGGAGCT GCTGGAGCTG
1981 GACAAGTGGG CCAGCCTGTG GAACTGGTTC AGCATCACCC AGTGGCTGTG GTACATCAAG
2041 ATCTTCATCA TGATCATCGG CGGCCTGATC GGCCTGCGCA TCGTGTTCGC CGTGCTGAGC
2101 CTGGTGAACC GCGTGCGCCA GGGCTACAGC CCCCTGAGCT TCCAGACCCT GCTGCCCGCC
2161 CCCCGCGGCC CCGACCGCCC CGAGGGCACC GAGGAGGAGG GCGGCGAGCG CGGCCGCGAC
2221 CGCAGCGTGC GCCTGCTGAA CGGCTTCAGC GCCCTGATCT GGGACGACCT GCGCAGCCTG
2281 TGCCTGTTCA GCTACCACCG CCTGCGCGAC CTGATCCTGA TCGCCGTGCG CATCGTGGAG
2341 CTGCTGGGCC GCCGCGGCTG GGACATCCTG AAGTACCTGT GGAACCTGCT GCAGTACTGG
2401 AGCCAGGAGC TGCGCAACAG CGCCAGCAGC CTGTTCGACG CCATCGCCAT CGCCGTGGCC
2461 GAGGGCACCG ACCGCGTGAT CGAGATCATC CAGCGCGCCT GCCGCGCCGT GCTGAACATC
2521 CCCCGCCGCA TCCGCCAGGG CCTGGAGCGC AGCCTGCTGT AA
```

FIGURE 17

gp160mod.93IN101

```
   1 ATGCGCGTGC GCGGCACCCT GCGCAACTAC CAGCAGTGGT GGATCTGGGG CGTGCTGGGC
  61 TTCTGGATGC TGATGATCTG CAACGGCGGC GGCAACCTGT GGGTGACCGT GTACTACGGC
 121 GTGCCCGTGT GGAAGGAGGC CAAGACCACC CTGCTGTGCG CCAGCGACGC CAAGGCCTAC
 181 GAGCGCGAGG TGCACAACGT GTGGGCCACC CACGCCTGCG TGCCCACCGA CCCCAACCCC
 241 CAGGAGATCG TGCTGGGCAA CGTGACCGAG AACTTCAACA TGTGGAAGAA CGACATGGTG
 301 GACCAGATGC ACGAGGACGT GATCAGCCTG TGGGACCAGA GCCTGAAGCC CTGCGTGAAG
 361 CTGACCCCCC TGTGCGTGAC CCTGGAGTGC CGCAACGTGA GCCGCAACGT GAGCAGCTAC
 421 AACACCTACA ACGGCAGCGT GGAGGAGATC AAGAACTGCA GCTTCAACGC CACCCCCGAG
 481 GTGCGCGACC GCAAGCAGCG CATGTACGCC CTGTTCTACG GCCTGGACAT CGTGCCCCTG
 541 AACAAGAAGA ACAGCAGCGA GAACAGCAGC GAGTACCGCC TGATCAACTG CAACACCAGC
 601 GCCATCACCC AGGCCTGCCC CAAGGTGACC TTCGACCCCA TCCCCATCCA CTACTGCGCC
 661 CCCGCCGGCT ACGCCATCCT GAAGTGCAAC AACAAGACCT TCAACGGCAC CGGCCCCTGC
 721 AACAACGTGA GCACCGTGCA GTGCACCCAC GGCATCAAGC CCGTGGTGAG CACCCAGCTG
 781 CTGCTGAACG GCAGCCTGGC CGAGGGCGAG ATCATCATCC GCAGCGAGAA CCTGACCAAC
 841 AACGTGAAGA CCATCATCGT GCACCTGAAC CAGAGCGTGG AGATCGTGTG CACCCGCCCC
 901 AACAACAACA CCCGCAAGAG CATCCGCATC GGCCCCGGCC AGACCTTCTA CGCCACCGGC
 961 GACATCATCG GCGACATCCG CCAGGCCCAC TGCAACATCA GCCGCGACAA GTGGAACGAG
1021 ACCCTGCAGC GCGTGGGCAA GAAGCTGGCC GAGCACTTCC ACAACAAGAC CATCAAGTTC
1081 GCCAGCAGCA GCGGCGGCGA CCTGGAGATC ACCACCCACA GCTTCAACTG CCGCGGCGAG
1141 TTCTTCTACT GCAACACCAG CGGCCTGTTC AACGGCACCT ACATGCCCAC CTACATGCCC
1201 ACCGGCACCG AGAGCAACAG CAACAGCACC ATCACCATCC CCTGCCGCAT CAAGCAGATC
1261 ATCAACATGT GGCAGGAGGT GGGCCGCGCC ATGTACGCCC CCCCCATCGC CGGCAACATC
1321 ACCTGCACCA GCAACATCAC CGGCCTGCTG CTGGTGCACG ACGGCGGCAT CAAGGAGAAC
1381 GACACCGAGA ACAAGACCGA GATCTTCCGC CCCGGCGGCG GCGACATGCG CGACAACTGG
1441 CGCAGCGAGC TGTACAAGTA CAAGGTGGTG GAGATCAAGC CCCTGGGCGT GGCCCCCACC
1501 GCCGCCAAGC GCCGCGTGGT GGAGCGCGAG AAGCGCGCCG TGGGCATCGG CGCCGTGTTC
1561 CTGGGCTTCC TGGGCGCCGC CGGCAGCACC ATGGGCGCCG CCAGCATCAC CCTGACCGCC
1621 CAGGCCCGCC AGCTGCTGAG CGGCATCGTG CAGCAGCAGA GCAACCTGCT GCGCGCCATC
1681 GAGGCCCAGC AGCACCTGCT GCAGCTGACC GTGTGGGGCA TCAAGCAGCT GCAGACCCGC
1741 GTGCTGGCCA TCGAGCGCTA CCTGAAGGAC CAGCAGCTGC TGGGCATCTG GGGCTGCAGC
1801 GGCAAGCTGA TCTGCACCAC CGCCGTGCCC TGGAACAGCA GCTGGAGCAA CAAGACCCAG
1861 AGCGAGATCT GGAACAACAT GACCTGGATG CAGTGGGACC GCGAGGTGAG CAACTACACC
1921 AACATCATCT ACAGCCTGCT GGAGGAGAGC CAGAACCAGC AGGAGAAGAA CGAGAAGGAC
1981 CTGCTGGCCC TGGACAGCTG GAAGAACCTG TGGAGCTGGT TCGACATCAC CAACTGGCTG
2041 TGGTACATCA AGATCTTCAT CATGATCGTG GGCGGCCTGA TCGGCCTGCG CATCATCTTC
2101 GCCGTGCTGA GCATCGTGAA CCGCGTGCGC CAGGGCTACA GCCCCCTGAG CTTCCAGACC
2161 CTGACCCCCA ACCCCCGCGG CCCCGACCGC CTGGGCCGCA TCGAGGAGGA GGGCGGCGAG
2221 CAGGACAAGG ACCGCAGCAT CCGCCTGGTG AACGGCTTCC TGGCCCTGGC CTGGGACGAC
2281 CTGCGCAACC TGTGCCTGTT CAGCTACCAC CGCCTGCGCG ACTTCATCAG CGTGGCCGCC
2341 CGCGTGGTGG AGCTGCTGGG CCGCAGCAGC TGGGAGGCCC TGAAGTACCT GGGCAGCCTG
2401 GTGCAGTACT GGGGCCTGGA GCTGAAGAAG AGCGCCATCA GCCTGTTCGA CAGCATCGCC
2461 ATCGTGGTGG CCGAGGGCAC CGACCGCATC ATCGAGCTGG TGCAGGGCTT CTGCGCGCCC
2521 ATCCGCAACA TCCCCACCCG CATCCGCCAG GGCTTCGAGG CCGCCCTGCA GTAA
```

FIGURE 18

gp160mod.cm235.V3con

```
   1 ATGGATGCAA TGAAGAGAGG GCTCTGCTGT GTGCTGCTGC TGTGTGGAGC AGTCTTCGTT
  61 TCGCCCAGCG CTAGCAACAA CCTGTGGGTG ACCGTGTACT ACGGCGTGCC CGTGTGGCGC
 121 GACGCCGACA CCACCCTGTT CTGCGCCAGC GACGCCAAGG CCCACGAGAC CGAGGTGCAC
 181 AACGTGTGGG CCACCCACGC CTGCGTGCCC ACCGACCCCA ACCCCCAGGA GATCCACCTG
 241 GAGAACGTGA CCGAGAACTT CAACATGTGG AAGAACAACA TGGTGGAGCA GATGCAGGAG
 301 GACGTGATCA GCCTGTGGGA CCAGAGCCTG AAGCCCTGCG TGAAGCTGAC CCCCCTGTGC
 361 GTGACCCTGA ACTGCACCAA CGCCAAGCTG ACCAACGTGA ACAACATCAC CAGCGTGAGC
 421 AACACCATCG GCAACATCAC CGACGAGGTG CGCAACTGCA GCTTCAACAT GACCACCGAG
 481 CTGCGCGACA AGAAGCAGAA GGTGCACGCC CTGTTCTACA AGCTGGACAT CGTGCCCATC
 541 GAGGACAACA AGACCAGCAG CGAGTACCGC CTGATCAACT GCAACACCAG CGTGATCAAG
 601 CAGGCCTGCC CCAAGATCAG CTTCGACCCC ATCCCCATCC ACTACTGCAC CCCGCCGGC
 661 TACGCCATCC TGAAGTGCAA CGACAAGAAC TTCAACGGCA CCGGCCCCTG CAAGAACGTG
 721 AGCAGCGTGC AGTGCACCCA CGGCATCAAG CCCGTGGTGA GCACCCAGCT GCTGCTGAAC
 781 GGCAGCCTGG CCGAGGAGGA GATCATCATC CGCAGCGAGA ACCTGACCAA CAACGCCAAG
 841 ACCATCATCG TGCACCTGAA CAAGAGCGTG GAGATCAACT GCACCCGCCC CAGCAACAAC
 901 ACCCGCACCA GCATCACCAT CGGCCCCGGC CAGGTGTTCT ACCGCACCGG CGACATCATC
 961 GGCGACATCC GCAAGGCCTA CTGCGAGATC AACGGCACCA AGTGGAACGA GGTGCTGACC
1021 CAGGTGACCG AGAAGCTGAA GGAGCACTTC AACAACAAGA CCATCATCTT CCAGCCCCCC
1081 AGCGGCGGCG ACCTGGAGAT CACCATGCAC CACTTCAACT GCCGCGGCGA GTTCTTCTAC
1141 TGCAACACCA CCCGCCTGTT CAACAACACC TGCATCGAGA ACGGCACCAT GGGCGGCTGC
1201 AACGGCACCA TCATCCTGCC CTGCAAGATC AAGCAGATCA TCAACATGTG GCAGGGCGCC
1261 GGCCAGGCCA TGTACGCCCC CCCCATCAGC GGCCGCATCA ACTGCGTGAG CAACATCACC
1321 GGCATCCTGC TGACCCGCGA CGGCGGCGCC ATCAACACCA CCAACGAGAC CTTCCGCCCC
1381 GGCGGCGGCA ACATCAAGGA CAACTGGCGC AGCGAGCTGT ACAAGTACAA GGTGGTGCAG
1441 ATCGAGCCCC TGGGCATCGC CCCCACCCGC GCCAAGCGCC GCGTGGTGGA GCGCGAGAAG
1501 CGCGCCGTGG GCATCGGCGC CATGATCTTC GGCTTCCTGG GCGCCGCCGG CAGCACCATG
1561 GGCGCCGCCA GCATCACCCT GACCGTGCAG GCCCGCCAGC TGCTGAGCGG CATCGTGCAG
1621 CAGCAGAGCA ACCTGCTGCG CGCCATCGAG GCCCAGCAGC ACCTGCTGCA GCTGACCGTG
1681 TGGGGCATCA AGCAGCTGCA GGCCCGCGTG CTGGCCGTGG AGCGCTACCT GAAGGACCAG
1741 AAGTTCCTGG GCCTGTGGGG CTGCAGCGGC AAGATCATCT GCACCACCGC CGTGCCCTGG
1801 AACAGCACCT GGAGCAACCG CAGCTACGAG GAGATCTGGA ACAACATGAC CTGGATCGAG
1861 TGGGAGCGCG AGATCAGCAA CTACACCAAC CAGATCTACG AGATCCTGAC CGAGAGCCAG
1921 AACCAGCAGG ACCGCAACGA GAAGGACCTG CTGGAGCTGG ACAAGTGGGC CAGCCTGTGG
1981 AACTGGTTCG ACATCACCAA GTGGCTGTGG TACATCAAGA TCTTCATCAT GATCATCGGC
2041 GGCCTGATCG GCCTGCGCAT CATCTTCGCC GTGCTGAGCA TCGTGAACCG CGTGCGCCAG
2101 GGCTACAGCC CCCTGAGCTT CCAGACCCCC TTCCACCACC AGCGCGAGCC CGACCGCGAG
2161 GAGCGCATCG AGGAGGGCGG CGGCGAGCAG GGCCGCGACC GCAGCGTGCG CCTGGTGAGC
2221 GGCTTCCTGG CCCTGGCCTG GGACGACCTG CGCAGCCTGT GCCTGTTCAG CTACCACCGC
2281 CTGCGCGACT TCATCCTGAT CGCCGCCCGC ACCGTGAAGC TGCTGGGCCG CAGCAGCCTG
2341 AAGGGCCTGC GCCGCGGCTG GGAGGGCCTG AAGTACCTGG GCAACCTGCT GCTGTACTGG
2401 GGCCAGGAGC TGAAGATCAG CGCCATCAGC CTGCTGGACG CCACCGCCAT CATCGTGGCC
2461 GGCTGGACCG ACCGCGTGAT CGAGGTGGCC CAGGGCGCCT GGCGCGCCAT CCTGCACATC
2521 CCCGCCGCA TCCGCCAGGG CCTGGAGCGC ACCCTGCTGT AA
```

FIGURE 19
gp160partialmod.cm235.V3 con

```
   1 ATGGATGCAA TGAAGAGAGG GCTCTGCTGT GTGCTGCTGC TGTGTGGAGC AGTCTTCGTT
  61 TCGCCCAGCG CTAGCAACAA CTTGTGGGTT ACAGTTTATT ATGGGGTTCC TGTGTGGAGA
 121 GATGCAGATA CCACCCTATT TTGTGCATCA GATGCCAAAG CACATGAGAC AGAAGTGCAC
 181 AATGTCTGGG CCACACATGC CTGTGTACCC ACAGACCCCA ACCCACAAGA AATACACCTG
 241 GAAAATGTAA CAGAAAATTT TAACATGTGG AAAAATAACA TGGTAGAGCA GATGCAGGAG
 301 GATGTAATCA GTTTATGGGA TCAAAGTCTA AAGCCATGTG TAAAGTTAAC TCCTCTCTGC
 361 GTTACTTTAA ATTGTACCAA TGCTAAGTTG ACCAATGTCA ATAACATAAC CAGTGTCTCT
 421 AACACAATAG GAAATATAAC AGATGAAGTA AGAAACTGTT CTTTTAATAT GACCACAGAA
 481 CTAAGAGATA AGAAGCAGAA GGTCCATGCA CTTTTTTATA AGCTTGATAT AGTACCAATT
 541 GAAGATAATA AGACTAGTAG TGAGTATAGG TTAATAAATT GTAATACTTC AGTCATTAAG
 601 CAGGCTTGTC CAAAGATATC CTTTGATCCA ATTCCTATAC ATTATTGTAC TCCAGCTGGT
 661 TATGCGATTT TAAAGTGTAA TGATAAGAAT TTCAATGGGA CAGGGCCATG TAAAAATGTC
 721 AGCTCAGTAC AATGCACACA TGGAATTAAG CCAGTGGTAT CAACTCAATT GCTGTTAAAT
 781 GGCAGTCTAG CAGAAGAAGA GATAATAATC AGATCTGAAA ATCTCACAAA CAATGCCAAA
 841 ACCATAATAG TGCACCTTAA TAAATCTGTA GAAATCAATT GTACCAGACC CTCCAACAAT
 901 ACAAGAACAA GTATAACTAT AGGACCAGGA CAAGTATTCT ATAGAACAGG AGACATAATA
 961 GGAGATATAA GAAAAGCATA TTGTGAGATT AATGGAACAA AATGGAATGA AGTTTTAACA
1021 CAGGTAACTG AAAAATTAAA AGAGCACTTT AATAATAAGA CAATAATCTT TCAACCACCC
1081 TCAGGAGGAG ATCTAGAAAT TACAATGCAT CATTTTAATT GTAGAGGGGA ATTTTTCTAT
1141 TGCAATACAA CACGACTGTT TAATAATACT TGCATAGAAA ATGGAACCAT GGGGGGGTGT
1201 AATGGCACTA TCATACTTCC ATGCAAGATA AAGCAAATTA TAAACATGTG GCAGGGAGCA
1261 GGACAAGCAA TGTATGCTCC TCCCATCAGT GGAAGAATTA ATTGTGTATC AAATATTACA
1321 GGAATACTAT TGACAAGAGA TGGTGGTGCT ATTAATACAA CTAATGAGAC CTTCCGCCCC
1381 GGCGGCGGCA ACATCAAGGA CAACTGGCGC AGCGAGCTGT ACAAGTACAA GGTGGTGCAG
1441 ATCGAGCCCC TGGGCATCGC CCCCACCCGC GCCAAGCGCC GCGTGGTGGA GCGCGAGAAG
1501 CGCGCCGTGG GCATCGGCGC CATGATCTTC GGCTTCCTGG GCGCCGCCGG CAGCACCATG
1561 GGCGCCGCCA GCATCACCCT GACCGTGCAG GCCCGCCAGC TGCTGAGCGG CATCGTGCAG
1621 CAGCAGAGCA ACCTGCTGCG CGCCATCGAG GCCCAGCAGC ACCTGCTGCA GCTGACCGTG
1681 TGGGGCATCA AGCAGCTGCA GGCCCGCGTG CTGGCCGTGG AGCGCTACCT GAAGGACCAG
1741 AAGTTCCTGG GCCTGTGGGG CTGCAGCGGC AAGATCATCT GCACCACCGC CGTGCCCTGG
1801 AACAGCACCT GGAGCAACCG CAGCTACGAG GAGATCTGGA ACAACATGAC CTGGATCGAG
1861 TGGGAGCGCG AGATCAGCAA CTACACCAAC CAGATCTACG AGATCCTGAC CGAGAGCCAG
1921 AACCAGCAGG ACCGCAACGA GAAGGACCTG CTGGAGCTGG ACAAGTGGGC CAGCCTGTGG
1981 AACTGGTTCG ACATCACCAA GTGGCTGTGG TACATCAAAA TATTTATAAT GATAATAGGA
2041 GGTTTAATAG GTTTAAGGAT AATTTTTGCT GTGCTTTCTA TAGTGAATAG AGTTAGGCAG
2101 GGATACTCAC CTTTGTCTTT CCAGACCCCT TTCCATCATC AGAGGGAACC CGACAGATCC
2161 GAAAGAATCG AAGAAGGAGG TGGCGAGCAA GGCAGAGACA GATCCGTGCG ATTAGTGAGC
2221 GGATTCTTAG CTCTTGCGTG GGACGATCTA CGGAGCCTGT GCCTCTTCAG CTACCACCGC
2281 TTGAGAGACT TCATCTTGAT TGCAGCGAGG ACTGTGAAAC TTCTGGGACG CAGCAGTCTC
2341 AAGGGACTGA GACGGGGGTG GGAAGGTCTC AAATATCTGG GGAATCTTCT GTTATATTGG
2401 GGTCAGGAAC TAAAAATTAG CGCTATTTCT TTGCTTGATG CTACAGCAAT AATAGTAGCG
2461 GGGTGGACAG ATAGGGTTAT AGAAGTAGCA CAAGGAGCTT GGAGAGCCAT TCTCCACATA
2521 CCTAGGAGAA TCAGACAGGG CTTAGAAAGG ACTTTGCTAT AA
```

Serum Binding Antibody Titers to HIV-1_{SF162} Env Protein

	Replication-Competent Ad		Replication-Defective Ad	
Immunization	**10^7**	**10^8**	**10^8**	**10^9**
Post 1st Ad	20 20 20	114.11 390.91 33.31	276.00 N/A	46.11 72.06 29.51
Post 2nd Ad	2315.60 14380.44 372.87	4242.53 8251.33 2181.35	114 N/A	55.57 128.00 24.13
Post 1st gp140	41175.45 42411.99 39974.95	43589.41 51950.41 36574.05	906 N/A	2675.15 9448.33 757.43
Post 2nd gp140	19789.57 32906.06 11901.37	65799.55 68333.17 63359.86	14176 N/A	105578.03 208905.20 53357.79

FIGURE 20

FIGURE 21

Proliferative responses following Ad-HIV*env* recombinant priming and HIV_{SF162} oligomeric gp140ΔV2 boosting